Research

Survey Attributes, Development, Utilization, and Interpretation in Healthcare Research

John Ward, DC, MA, MS¹*

Address: ¹Associate Professor/Research Fellow, Department of Physiology and Chemistry, Texas Chiropractic College, Pasadena, TX, USA.

E-mail: John Ward, DC, MA, MS – Jward@txchiro.edu

*Corresponding author

Topics in Integrative Health Care 2015, Vol. 6(3)  ID: 6.3002

Published on December 31, 2015 | Link to Document on the Web

Abstract

Introduction: Surveys are important for the evaluation of healthcare data trends that affect patient populations.

Methods: This narrative review describes survey attributes, development, utilization, and interpretation from perspectives pertinent to researchers.

Results: Survey questions should be sensitive to cultural, psychological, and economic factors. Questions may be open-ended or closed-ended. They may be distributed in-person, through telephone interviews, by mail, online, or through direct physical distribution. Newly created surveys should use age-appropriate, education level-appropriate, culturally sensitive, and time-sensitive questions. Researchers will need to select an appropriate mechanism of quantifying and analyzing responses. Newly created surveys should be pilot-tested on small sample populations that are representative of the intended larger population. There is no set agreed upon way to report surveys, but the most common survey reporting deficiencies appear to be failing to: provide survey questions, report on validity or reliability of the instrument, provide the response rate, discuss how well the sample represents the entire population, and describe how missing data are handled.

Conclusions: Surveys are critical to discovering data trends and encouraging modifications in clinical treatment guidelines. The questions should focus on set domains, but be capable of accommodating all relevant answers. Survey questions should be tested to ensure they have reasonable validity for what they are trying to measure and that they are reliable if used multiple times with various populations.
Introduction

Surveys are important for research and evaluation of data trends within patient populations.\textsuperscript{1-4} The information they provide can lead to improvements in healthcare practice patterns through the development of treatment guidelines.\textsuperscript{5-9}

The three main categories of surveys are: descriptive, comparative, and predictive.\textsuperscript{10} Descriptive surveys provide characteristic information about a specific population. Comparative surveys are used to compare and contrast the attributes of two or more groups. Predictive surveys forecast the actions of a particular group of individuals based on their previous patterns in life.

Advantages of using surveys are that they can: 1) reach many respondents with minimal effort, and 2) process how multiple categorically different variables interact with one another.\textsuperscript{11-12} A disadvantage of surveys is that low response rates can threaten their external validity.\textsuperscript{11}

Methods

Topically, the section subheadings of this narrative review were designed to be similar to most of the individual titles of Fink’s nine-volume textbook series on developing surveys.\textsuperscript{10,13-17} The specific aims of this article are to describe the attributes of surveys, and how they are developed, distributed, and analyzed.

How should survey questions be phrased?

Survey questions should be sensitive to cultural,\textsuperscript{13,18-23} psychological,\textsuperscript{13} and economic factors.\textsuperscript{13} Ideally, questions will make sense to the respondent,\textsuperscript{13,24} be relevant to important time periods in his/her life, and avoid expression of bias.\textsuperscript{13} Loaded questions should be avoided.\textsuperscript{13,25} An example of a loaded question would be for a doctor to ask a patient with a general headache, “How long have you been having migraines?” This question is pressuring the patient to accept that they have a set type of headache even though they may have another type (i.e. cluster, tension, or sinus headache). If possible, researchers should adopt questions from other surveys that have already been validated.\textsuperscript{13}

Questions can be open-ended or closed-ended.\textsuperscript{12-13,25} Open-ended questions allow respondents to write their own answers. These can explain a respondent’s motivation when it is not clear to the researcher.\textsuperscript{13,25} A disadvantage of using open-ended questions is that the responses need to be coded by the researcher, which allows for subjective interpretation.\textsuperscript{25} Closed-ended questions offer predetermined answer choices and are faster for researchers to analyze. One disadvantage of closed-ended questions is that sometimes the most appropriate response to a question is missing from the given options for a participant.\textsuperscript{25}

Types of survey distribution

Surveys may be administered in-person, through telephone interviews, by mail, online, or by direct physical distribution to participants for return at a later date.
If surveys are administered in-person by multiple surveyors, then the surveyors should be trained to ask questions using a standard format or similar phrasing. For example, one researcher might ask a respondent, “are you married,” while a different researcher asks, “are you married or do you have a life partner.” If the primary purpose of the survey question was to see if the person lives alone then the second researcher would be asking a more informative question that would take into account same-sex couples. Quality checks of each surveyor should be performed to ensure standardized phrasing. One advantage of using in-person surveys is that questions can be asked by respondents to ensure they clearly understand the meaning of particular survey questions. There are two significant disadvantages of using in-person surveys. They have been found to result in more positive scores in some fields compared to other forms of survey administration. Additionally, studies have shown that research participants are less likely to report morbidity or socially inadequate behaviors during in-person surveys. As a result, human interaction during the administration of a survey can impact the findings produced.

Telephone surveys possess their own unique attributes, particularly when compared to mail-out surveys. Comparison studies between telephone surveys and mail-out surveys have shown that telephone surveys are less expensive, are more likely to be fully completed, and demonstrate an increased chance a person will choose to take the survey in the first place. A disadvantage of telephone surveys is that they typically involve random-digit-dialing (RDD) to landline telephone numbers and not cell phones. As a result, those potential respondents who only use cell phones, or who have unlisted landlines will be missed.

Mail surveys involve sending paper surveys directly to respondents. Two significant advantages of mail surveys are that they can contain lengthy questions that a participant may fill out at their convenience, and they can include pictures. Two significant disadvantages of mail surveys that have been reported are greater levels of missing data and higher numbers of inconsistent answers.

Online surveys are directly hosted on websites. This allows researchers to force participants to answer a given question before proceeding to the next question. There are several online survey platforms that can be used, including: REDCap, Survey Monkey, Zoomerang, Qualtrics, and QuesGen. Online surveys result in savings on postage and printing costs and can use skip questions. The latter refers to the display of subsequent questions based on initial responses. For example, only respondents who answer “female” to a question about gender will later see a series of female-only reproductive questions in their survey. Unfortunately, online surveys often have low response rates. Other disadvantages are a bias toward younger, more computer-literate participants and the exclusion of individuals who do not have access to a computer.

Handout surveys are distributed by healthcare providers directly in their office for later deposit into a drop box or to be returned via mail. An advantage of in-office surveys is that they are cheaper than mail-out and telephone surveys. One disadvantage is that in-office handout surveys have the potential for measurement bias if office gatekeepers distribute the survey to selected patients, rather than according to a predetermined random pattern.
Factors to consider when designing a new survey

Figure 1 lists seven questions that should be answered when designing a new survey. As survey questions are developed, researchers should look at existing surveys and attempt to emulate questions that are relevant to their topic. This provides a standard by which the results of the newly developed survey can be compared.

Fig.1 Questions to answer as a survey is being developed.
Questions to ask prior to developing a survey
1) What is the purpose?
2) Who are the respondents?
3) How will the survey be administered?
4) How will the response be tallied (e.g., a rating system, or categorically unique answers)
5) How much time will respondents be expected to spend on the survey?
6) Will the survey need to be translated to another language?
7) Will the respondents be anonymous?

Prior to creating a survey, it is important to understand the end purpose of the information that researchers are trying to gain. This requires investigators to be familiar with existing research on their topic and to develop a survey that helps identify relevant gaps in existing information. By identifying these gaps, future interventional studies can be developed to improve overall healthcare outcomes.

Questions must be developed with respondents in mind. The questions should be age-appropriate, education level-appropriate, culturally sensitive, and time-sensitive. Researchers need to make an effort to write questions that most respondents will be able to understand. Questions should not ask participants to recall information that happened too long ago, because this may increase the chance they will not remember the information accurately. Predetermined inclusion and exclusion criteria should also be established to ensure that the correct population is receiving the survey. If a population is not selected appropriately, the internal and external validity of the study may be affected.

Surveys can be administered in-person, through telephone interviews, mailed, online, or through in-office handouts. Each option has its own strengths and weaknesses, some of which have already been discussed.

Survey answers can be based on a rating scheme or they can involve nominal categories (e.g., patient race, religion). The Likert scale is an example of a rating scheme that uses a point value system. This scale can involve positively-scaled options or negatively-scaled options. The following example is of negatively-scaled Likert options: strongly agree (1), agree (2), neutral (3), disagree (4), and strongly disagree (5). With this form of scaling the more negative responses are worth a higher point value. When using Likert scales, some researchers suggest utilizing a combination of negatively-scaled and
positively-scaled questions, which is referred to as a balanced scale, to reduce the response set effect. This effect is seen when participants choose one given answer many times in a row (e.g., strongly disagree) because they feel it likely applies to most questions.\textsuperscript{55} Other less-common rating scales include: Rasch model scaling, interval-ratio scaling, and semantic differential scaling.\textsuperscript{12}

When surveys are developed, researchers should create domain-specific questions if possible that can be answered in a reasonable amount of time.\textsuperscript{56-57} For example, the Short Form (SF)-36 has a domain for physical health problems, emotional health problems, social activities, and other domains.\textsuperscript{58} If a survey includes too many questions, respondents may not want to complete the survey and may even answer questions inaccurately.

If a survey is intended to be distributed in multiple languages, then interpretation will be required. Ideally, researchers will translate the survey to the new language and then have it independently translated back to see if the content is being relayed properly.

Surveys can either be anonymous or not. The benefit of anonymous surveys is that they allow researchers to ask more personal questions. For example, a researcher could ask participants about their sexual history.

**Visual considerations in creating a survey**

Surveys typically should utilize size 10 font or greater.\textsuperscript{13-14} If researchers are anticipating that some of their survey respondents will have deficiencies with reading text (i.e., presbyopia), they can make the font even larger.\textsuperscript{25} When open-ended questions are used, the amount of open space left for responses is important. Studies have shown that the size of open space for writing on the survey is directly proportional to how much the respondents will write.\textsuperscript{59-60} Finally, research has demonstrated that shorter surveys have higher response rates.\textsuperscript{56-57}

**How to test a survey on a small population**

Prior to utilizing a new survey, it is ideal to test it using a small group of sample respondents. This can be used to determine whether any of the questions are misunderstood. Testing a survey out on a small group of respondents, individually, or with a focus group can also identify unforeseen issues that may arise later when the survey is fully administered to a large population.\textsuperscript{12,14,61} **Figure 2** illustrates some of the issues researchers should review when testing their survey.

**Fig. 2. Survey testing questions a researcher should answer.**

Questions that should be answered by survey testing

1) Are the survey questions appropriate for respondents?
2) Are any questions misleading?
3) If the survey is given through interview are the surveyors able to use the surveys appropriately?
4) Is the information obtained by the survey reliable?
5) Is the information obtained by the survey valid?
When pilot-testing a survey, researchers should choose a group of participants that is representative of the target population they ultimately want to sample. This can be done through probability sampling or non-probability sampling.\textsuperscript{12,14}

Probability sampling means that any person within a known target population of the study might be selected.\textsuperscript{12,14} This includes: simple random sampling, stratified random sampling, systematic sampling, and cluster sampling. Simple random sampling is typically performed when a list of participants in a given population is matched up against a predetermined random number list (e.g., recruit applicant #1, 5, 12, 14).\textsuperscript{12,14} Thus, participants out of the sample pool will be recruited at random. Stratified random sampling involves a population being divided into subgroups, termed strata, and then a random sample is taken from those subgroups.\textsuperscript{12,14,62} For example, a researcher may want to sample participants with low back pain as the main population. The researcher then further develops strata based on age groups (e.g., divide the population into: below 18, 19-30, 31-50 and 51+ years-of-age strata) and analyzes low back pain as it applies to the various age groups. Systematic sampling involves picking participants at set intervals. An example would be selecting every 5\textsuperscript{th} person out of the total population for the study.\textsuperscript{12,14} Finally, cluster sampling involves selecting participants within some form of cluster (e.g., participants from one physical therapy clinic, as opposed to all physical therapy clinics in a city).

Non-probability sampling means that only a particular set of individuals is chosen to participate in the survey.\textsuperscript{12,14} This can involve convenience sampling, snowball sampling, quota sampling, or focus groups.\textsuperscript{12,14} Convenience sampling utilizes individuals that are readily available for sampling based on opportunity. An example would be a classroom of student volunteers.\textsuperscript{12,14} Snowball sampling involves surveying participants and then asking them to recommend people they know in the same demographic which is studied.\textsuperscript{12,14,63} Quota sampling involves determining a set number of participants in given subgroups that need to be sampled and then recruiting that quantity.\textsuperscript{12,14} For example, a researcher could choose to sample 50 men and 50 women with whiplash. Finally, focus groups typically involve structured interactions between 10-20 participants during which survey questions can be reviewed.\textsuperscript{64-66}

When administering a survey, response rates may vary. There is no set standard response rate. Unsolicited mail-out surveys often can have response rates as low as 20%.\textsuperscript{13,17} Survey response rates may be increased by using incentives. Studies have shown that monetary or gift incentives significantly increase survey response rate.\textsuperscript{12,56,67-74}

**Measurement of survey reliability and validity**

Psychometrics represents the study of the design, administration, and interpretation of quantitative assessments.\textsuperscript{16} It often involves the evaluation of reliability and validity. The quantification of these attributes is important for the assessment of the quality of a survey.

Reliability refers to the reproducibility of the data from a survey instrument. Test-retest reliability is utilized when set respondents repeat the survey at two different time points.\textsuperscript{16,75} Correlation coefficients (e.g., $r$ values) can be developed to compare the responses of a group of respondents at two different time points. Values for $r$ at or above 0.7 are considered good.\textsuperscript{16} The practice effect, or the bias from remembering one’s previous answers to survey questions, can be treated as a covariate when the same group is surveyed twice in close temporal succession. A covariate is a side variable that is not directly controlled which can impact the main dependent variable studied. To solidify this point, assume a
researcher is tracking athlete caloric intake based on different age groups. If the researcher is measuring data throughout the entire year the researcher’s findings could be impacted during the winter holidays (i.e. Thanksgiving), thus “season” could act as a covariate. The practice effect can be handled by changing the order of the survey questions or by altering a few words in each question without changing the context of the questions.\textsuperscript{16}

Validity refers to how well a survey measures what it is intended to measure. The different kinds of validity include face, content, criterion, and construct validity.\textsuperscript{16} Face validity involves showing a survey to untrained individuals and asking them if the survey appears to reasonably address its topic.\textsuperscript{16,25} This is the least stringent method of validity.\textsuperscript{16} Face validity can be measured by providing respondents with a survey and then asking them afterward if they feel the survey focused on a particular topic. If the survey did not focus on the intended topic, then the researchers could improve the questions. Content validity involves having a set of one or more subject matter experts review a survey to determine if it only includes what it should.\textsuperscript{16} This form of validity involves multiple experts rating how much they agree with a given statement in a survey being essential to the intended topic domain. For example, it could involve raters reviewing a statement and deciding that it is ‘essential’, ‘useful, but not essential’, or ‘not essential’. Criterion validity provides quantitative evidence as to how well one survey compares to another.\textsuperscript{16} For example, a new low back pain survey could be given to low back pain patients along with the more established Roland-Morris Low Back Pain and Disability Questionnaire.\textsuperscript{76} Then the researcher could compare the scores on the new survey against the more established survey. Construct validity is the most valuable form of validity, but it often takes years to measure.\textsuperscript{16} It is essentially a measure of the degree to which a survey measures what it claims to measure after being analyzed with multiple populations over several years.

How to report survey data

At a minimum, the authors of a survey journal article should introduce the significance of their topic, describe the methods and results of their survey, review their conclusions, and discuss the implications, recommendations, and future directions of research.\textsuperscript{17,77} Bennet et al reviewed 117 surveys and concluded that there is no standard way of reporting survey research.\textsuperscript{76} However, the most common weaknesses they observed were failing to: provide survey questions (35%), report on validity or reliability of the instrument (19%), provide the response rate (25%), discuss how well the sample represents the entire population (11%), and describe how missing data were handled (11%).\textsuperscript{77}

Survey data can be reported graphically by utilizing percentages, pie charts, bar graphs, line charts, shaded regional maps, and other means. Data tables generally should report sample size (n) and overall percentage for each population sampled, per question. Finally, researchers should avoid scaling the Y-axis to prevent readers from misinterpreting bar graph data, if utilized.\textsuperscript{17} Assume a researcher wants to compare two groups of students with a survey where scores are scaled from 0 to 100%. Group A scores a 75% average, while group B has a 90% average. If a researcher starts the Y-axis at 70% for their graph it will make the difference in performance between group A and B appear large (Figure 3). However, if the researcher starts the Y-axis at 0, the true lowest axis for the data, the relationship between group A and B will be more accurately demonstrated.
Conclusions

Surveys are valuable for discovering data trends in populations. Many factors must be considered when developing and implementing new surveys to include an understanding of cultural, psychological and economic factors that may impact the population studied. Based on these observations, future interventions can be developed to improve healthcare outcomes. An example of this occurring would be how the Institute for Clinical Systems Improvement (ICSI) developed adult acute and subacute low back pain health care guidelines based on reviewing several health surveys.  

Acknowledgement

The author of this manuscript would like to thank Claire Noll, M.L.I.S. for assistance with editing.

References


33. Pederson L, Baskerville J, Ashley M, Lefcoe N. Comparison of mail questionnaire and telephone
interview as data gathering strategies in a survey of attitudes toward restrictions on cigarette smoking. 


