

Survey Science in Pediatric Emergency Medicine

Mark R. Zonfrillo, MD, MSCE*† and Douglas J. Wiebe, PhD‡

Abstract: Surveys conveniently acquire and summarize valuable information from a target population. The specific aims, design, target sample, mode of distribution, data analysis, and inherent limitations of the survey methodology should be carefully considered to maximize the validity of the results. This review provides guidance on the methods and standards necessary to complete sound survey science.

Key Words: survey, questionnaire

(*Pediatr Emer Care* 2011;27: 443–448)

BACKGROUND

Surveys are questionnaires distributed by the government, private corporations, the media, scientists, and a variety of other groups to obtain information from a target population. Surveys are often used by pediatric emergency medicine (PEM) researchers to gain an understanding of patient perspectives or to frame attitudes toward clinical practice. As a subspecialty, PEM has grown substantially since its inception more than 30 years ago, and surveys have become increasingly more popular within PEM research. Of all PEM abstracts presented at national conferences between 1987 and 1999, the percentage of descriptive surveys increased from 9% to 16%.¹ Although surveys appear enticingly convenient and inexpensive, they must be carefully designed to accurately meet the objectives of the investigators (Fig. 1). When conducted properly, surveys can be useful tools for the PEM community, and investigators should be familiar with recommendations and guidelines for best survey practices. This review aims to explore the science behind survey design, administration, and analysis. There is a focus on written and electronic surveys and on issues germane to PEM-based content.

LITERATURE REVIEW

A careful review of the literature is important to identify surveys in similar subject areas or for similar populations. Expanding on existing work can be useful, especially if a previously conducted survey instrument was well designed and validated. Particular attention should be paid to the studies' methods to recognize sound survey science. Contacting authors and other experts about their challenges with previous surveys can be helpful to identify and overcome methodological limitations. Information from the literature can be augmented with focus groups to delineate the broad objectives and item generation.

SPECIFIC AIMS

Before designing a survey, investigators must explicitly identify the project's specific aims, which serve as a foundation

for methodology and statistical analyses. Although survey items can assay a variety of information (eg, attributes, knowledge, attitudes, experiences), the specific aims should focus on the most important objectives that are identified by the investigators. This is especially crucial when developing surveys because there may be a tendency to include items that may not fulfill the primary objectives of the questionnaire. Longer surveys with unnecessary items could potentially decrease the response rate and yield excessive data. Therefore, survey design should include *a priori* description of the proposed analyses to include only the relevant survey items.

DEFINING THE TARGET POPULATION

The target population should be clearly delineated when developing the specific aims of the survey. Describing the target population has implications for the survey methodology, potential bias from respondents and in responses, and generalizability of the survey results. Investigators should describe the rationale for choosing their target population and aim to obtain a sample representative of the entire group. Investigators should describe the rationale for choosing their target population and aim to obtain a sample representative of the entire group, although this may be infeasible for a variety of reasons. In such instances, the differences expected between the sample that was obtained and the whole target population should be described in the manuscript.

SURVEY DESIGN

Modality

Surveys can be delivered via postal mail, by telephone, electronically, or in person. The choice of modality depends on the survey questions, the target population, and available resources. Most surveys use only 1 modality, although there are a few advantages of multimodal surveys, including increasing response rate and minimizing bias linked to a certain mode.² Although in-person interviews are useful for longer surveys and for obtaining more qualitative data, they are time consuming, costly, and not anonymous.³ Telephone interviews were a mainstay for many types of surveys but are increasingly more difficult to use in the era of mobile phones, "do not call" registries, calling number identification, and voice mail.³ Mailed surveys are convenient and inexpensive, but they require more work for printing and distribution and more active effort by the participant that can delay data collection.³ Electronic-based surveys, including computer- and e-mail-based questionnaires, are inexpensive, efficient, and may minimize transcription error.² However, they may be more subject to selection bias and lower response rates.³ Selection bias can occur from variability in access to or interest in e-mail or from individuals only responding to surveys that are of interest to them. Lower response rates may occur from e-mail being filtered as "spam" or respondents being unfamiliar with the sender. In addition, because Internet-based surveys are unmonitored, respondents may not complete the survey, particularly for longer questionnaires.

Sampling

There are a number of different sampling methods, the choice of which is determined by the specific aims of the survey

From the *Division of Emergency Medicine, The Children's Hospital of Philadelphia; Departments of †Pediatrics, and ‡Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA.

Reprints: Mark R. Zonfrillo, MD, MSCE, Division of Emergency Medicine, The Children's Hospital of Philadelphia, 34th St and Civic Center Boulevard, Philadelphia, PA 19104 (e-mail: zonfrillo@email.chop.edu).

The authors have no disclosures.

Copyright © 2011 by Lippincott Williams & Wilkins

ISSN: 0749-5161

and the resources available. Nonprobability sampling (ie, purposeful sampling that does not use randomization) may introduce biases in the data, most commonly selection bias, thereby limiting the conclusions that can be made.^{4,5} Probability sampling (ie, random sampling in which the entire population has a chance of being selected) can be simple, systematic, stratified, or clustered.⁴ Since the generalizability of survey data is determined by comparing the characteristics of respondents and nonrespondents, the random nature of probability sampling is preferable. When using a nonprobability sample, the population for which the data apply should be clearly stated in the methods and limitations. The sampling method influences the degree to which there is sampling error. This type of error stems from sample-to-sample variability: if separate random samples were chosen from the same group, the samples would be inherently different. The use of sampling weights corrects for the disproportionality of a sample with respect to the target population and enables population estimates to be derived from a survey.⁶ Guidance from a survey methodologist may be sought to determine if and how weights should be generated and incorporated during analysis.

Sample Size

Although some surveys are designed to reach all individuals in a population (eg, all pediatric emergency medicine physicians in the United States), it is typical to survey a sample. Survey sample size determination can be challenging. There are a few approaches that are commonly, but improperly, used.² The first is to specify the proportion of the target population (eg, 5%) to be included in the sample. Unfortunately, defining a proportion of the population does not necessarily prevent sampling error; the sample chosen may not be similar to the population as a whole. One exception to this is that when the target population is smaller, a sample of 10% or more of the population can reduce sampling error.² In general, however, a larger sample does not guarantee that it is representative, and the methods used to obtain a nonbiased sample are more important than the absolute sample size.

Another incorrect method is to use a standardized sample size for a particular population. For example, it is incorrect to say that 500 participants are adequate for a community sample, and 1500 respondents are adequate for a national sample.² Although

previous studies may serve as guidance to help gauge a reasonable sample size, their validity may not be apparent and should be carefully considered. Ultimately, investigators should individualize their sample size calculation based on the goals of their research and not by an absolute number for a certain group.²

A third common method of calculating sample size is to determine, a priori, the precision of a point estimate and/or allowable margin of error of a survey item.² Although this can be a useful starting point for sample size estimation, investigators must understand the inherent limitations of this method. There are often multiple questions and objectives for a survey, and these estimates may have a similar relevance for the investigators. For example, the primary aim of the study may be met by a much smaller sample size than other important secondary aims. Another limitation to this method is that it only accounts for the margin of error in sampling, whereas surveys have multiple sources of error.

Sample size calculations should start with a review of the study's specific aims and proposed statistical analyses. Investigators should consider the anticipated effect size and the anticipated survey response rate based on previous surveys and pilot data. Attention should be focused on the smallest subgroups and the minimum sample size tolerable to base conclusions on the data.² The investigator can then perform traditional sample size calculations with relevant equations or charts, depending on the study design (eg, descriptive, experimental) and nature of estimates (eg, proportions, means) to be compared.^{4,7} Given the complexity of sample size calculation, assistance from biostatisticians and experts in the subject of the survey should be considered. In addition, investigators must continuously reevaluate the methods of data collection because quality is far more valuable than quantity when surveying a sample.

Introduction

A cover page or introduction serves to introduce the investigators, the objective of the survey, and any relevant ethical considerations. The survey authors should briefly describe any scientific development, including pilot testing. They should include a time estimate needed to complete the survey and advanced appreciation for participation. The introduction, like the rest of the survey, should be clear, concise, and free of spelling and grammatical errors.

Measurements

The 4 most common levels of measurement in surveys are nominal data (eg, male/female, yes/no, and other categorical data that cannot be ranked), ordinal data (eg, always/sometimes/never, as seen in Likert scales⁸ and other ranked data), interval data (eg, dates, temperature in Fahrenheit, and other data that are ranked, numerical values with an arbitrary zero), and ratio data (eg, height, age, and other data with an absolute zero and for which the proportion of values determines relevance).⁴ There is debate over whether items with ordinal data should have an even or odd number of choices, and regarding the ideal number of choices. An even number of choices forces the respondent to commit in 1 general direction.³ This can be helpful to categorize or dichotomize survey data, especially because a small response rate may limit the power of the analyses. An odd number of questions will give respondents a neutral option, which can also be useful depending on the content of the item. The number of choices for Likert scales is typically 4 to 5 and should not exceed 6 to 7. For all types of questions, it is important that the choices are omniscient (ie, that all possible answers are listed),⁹ including choices that are neutral (eg, "neither like nor dislike"),

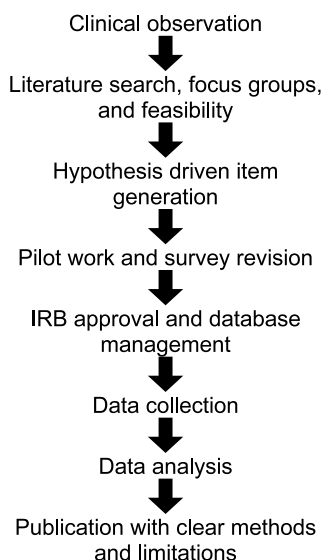


FIGURE 1. An outline of survey development and implementation.

unknown (eg, “do not know”), not applicable (eg, “does not apply”), and open ended (eg, “Other, please specify”).

Open-Ended and Closed Questions

Questions for nominal data can be open or closed, with the latter typically being more common and often more amenable to quantitative analyses. The researcher should carefully consider the reason for, the design of, and any potential answers to open-ended questions to maximize the usefulness of the responses. These types of questions can be successfully used as a source of qualitative data, revealing issues that may not have been conceived by the investigators and aiding future hypothesis generation. One of the most valuable open-ended questions is an item at the end of the survey that solicits other comments about the questionnaire. With the exception of these carefully defined intentions, questions should be closed ended. Closed questions should have all possible choices present, should have information as specific as possible, and should only include ranges of data that have some predetermined relevance.

Question-and-Answer Order

The ordering of questions can have implications for respondents' answers. There are advantages to various types of question ordering, which can be particularly easy to implement with electronic surveys. It may be useful to filter out questions that do not apply to respondents based on previous responses with the use of “skip logic,” also known as conditional branching. This technology can be easily implemented in electronic-based questionnaires to customize the survey and omit unnecessary items. It can also be used to determine respondent's eligibility, in which case the entire survey is skipped with redirection to a “thank you” page. Question order should also be carefully considered for a section of related items as order effects can be observed.^{10–12} Sections of questions and/or responses of electronic surveys can be configured in random or rotated order, which can be used to avoid bias.¹³ However, not all answer choices or questions should be randomized. For example, ordinal items should always be listed consistently. If there are items that could influence the responses of other questions, they should be put at the end of the survey. The content could serve as a cue for a knowledge item, or as an emotional trigger for an attitude item, and lead to biased responses. Electronic surveys can prohibit respondents from changing responses to earlier questions after they view subsequent survey items that could influence their answers. Demographic information and objectionable questions should also be reserved for the end of the survey.⁴

Formatting and Length

Survey length is closely related to the primary and secondary aims of the survey. “Less is more” is an important overarching rule, both to maximize response rates and to focus the survey. Every question should have a purpose and should be carefully constructed based on *a priori* hypotheses. Software programs used to create electronic surveys will typically have formatting that facilitates ease of reading for the respondent. Electronic surveys should include multiple screens to separate content area and provide small breaks for the reader. Fields of view vary across computer types, so the layout of a survey should be viewed on various computers during its development. Entry boxes such as radio buttons (ie, option buttons) should be used whenever possible to provide consistent, accurate, and complete responses to questions that have choices.¹⁰ Because there is conflicting evidence about how the size of a response field affects the written or typed answers in open-ended questions,¹⁰ questionnaires with response fields should be discussed in focus groups or tested in

pilot data. Survey progress indicators may keep subjects motivated to complete the survey but may be discouraging for longer surveys if the progress is minimal on each page.¹⁰ Accentuation with bold, underlined, and colored text should be used judiciously to highlight areas of the survey. Survey item choices should ideally be 1 per line, and all similar items should be formatted consistently throughout the survey. Electronic drop-down lists can reduce clutter but can bias the results if respondents have to scroll through a number of items.^{3,13} Forced choice programming should be used appropriately; respondents should only have 1 choice for certain items, but multiple choices should be allowed for “Choose all that apply” questions. The misuse of forced choice programming can make analysis extremely difficult and potentially invalidate some questions.

MAXIMIZING VALID RESPONSES

Respondents may not accurately respond to a question because they do not understand the question, they do not know or cannot recall the answer, they misestimate the answer, they do not follow the survey instructions, or they intentionally do not want to report the accurate answer.⁵ The use of clearly defined terms and survey piloting can maximize question comprehension. It is difficult to maximize recall of information that a respondent cannot remember, especially for minor or more distant events and information. When asked about their behavior, respondents will often estimate answers based on patterns instead of recalling actual events. Questions can be made easier by adjusting their objectives and assessing only necessary information. Finally, social desirability can contribute to inaccurate responses.² Surveys that emphasize anonymity or confidentiality and questions that reserve judgment will have higher response rates and more valid answers. Preceding a question with a statement that normalizes the topic at issue may foster honest responses.

DATABASE MANAGEMENT

A database should be established before disseminating the survey. It is important to create numeric codes for the survey data and to be consistent with the code for questions with similar responses. There should be codes for questions that are unanswerable (eg, “not applicable,” “do not know,” “other”) to minimize data entry error.² For open-ended or “free-text” questions, there should be a consistent method of capturing the data so that it may be analyzed manually. For many commercially available online survey tools, the data can be easily exported in a comma-separated values or another database-compatible format. This can reduce coding errors by eliminating transcription of paper-based data.¹³ Databases should be tested using mock or pilot data to identify any potential difficulties or inconsistencies with data entry.

SURVEY DESIGN AND PILOTING

Surveys should ideally be designed by a research focus group that includes members of the target population and then pilot tested in a similar group of individuals. This will ensure that the subject can navigate the survey easily, understand the content, and be able to accurately answer the items. It will also enhance the reliability of the survey by ensuring that the questions have the same meaning for all respondents. When pilot testing, the respondents should complete the survey as designed to determine the length of time for completion. In addition, test respondents should give specific written or oral feedback for the survey items. Electronic surveys should be piloted in their native form to identify any problems with hypertext links, survey page navigation, and skip logic.²

INSTITUTIONAL REVIEW BOARD

All survey activities should have approval from the local institutional review board before dissemination. Many surveys are minimal risk, especially when responses are anonymous, and will typically be eligible for expedited review or exemption. However, it is imperative that their content be formally reviewed because all research on human subjects needs formal regulation and evaluation.¹⁴ Full disclosure of the research activities should precede the survey, including the organization, investigator names, any sponsorship, goals of the survey, any risks or benefits to the participant, a statement of confidentiality and/or anonymity, a statement about voluntary participation, and contact information if the respondent has any questions.² This information allows for implied consent for institutional review board purposes. There are some situations where written consent may be necessary, especially when sensitive information is collected or if the subjects are in a vulnerable population. Although remuneration may be helpful in increasing response rates, it should be modest enough as to not be persuasive or coercive.¹⁵

SUMMARIZING SURVEY DATA

Response Rates

The first step in summarizing survey data is calculating response rates. Every effort should be made to assess the percentage of nonresponders and to compare their characteristics with those of the responders. These 2 factors contribute to the overall potential bias in the sample. If a survey is sent to an electronic mailing list, the list administrator should have basic summary statistics of the group. Although nonresponse clearly can affect the estimates measured, it is difficult to predict when and to what degree this bias occurs.²

Proven efforts to increase response rates include recognizable survey administrators or sponsors, user-friendly surveys, incentivization, and reminders. Many health care–based surveys are from individuals or research groups, and collegiality may promote response rate. Surveys should be designed in a clear concise format that is free of spelling and grammatical errors. Modest payments can be effective in increasing survey response rates. For mail questionnaires, prepayment of a few dollars increases response rates,² although a promise of payments and prize lotteries is less successful. With the deluge of daily electronic and postal mail, reminders are essential to emphasize the importance of the questionnaire.

Data Analysis

Although a detailed description of data analysis is beyond the scope of this discussion, there are some general principles of data entry and coding that should be emphasized. As previously noted, establishing a database and codes for various data responses can minimize transcription errors. Independent checks of coders' work are useful to identify errors and imperfect coding systems.² The close regulation of data entry and coding is a critical step in maximizing the validity of the data.

Many surveys use basic descriptive analyses (eg, frequencies, ratios, means), which are often sufficient to meet the overall objectives. Categorical answers (eg, those of a Likert scale) can be consolidated into larger groups or dichotomized, especially when cross-tabulated with other variables. Multivariate analysis can be used to account for multiple variables but should be reserved for surveys with larger response rates to draw the most valid conclusions. Although it is more difficult to complete quantitative analyses for qualitative data, these responses can be presented in raw form, categorized into broad groups, and used for future further item generation.

DATA FALSIFICATION

Falsification of data and other scientific misconduct is reprehensible yet, unfortunately, it is prevalent. One recent systematic review and meta-analysis of survey data found that approximately 2% of investigators admitted to data fabrication, falsification or modification of data, and up to one third admitted to other questionable research practices.¹⁶ As with all research, surveys should be conducted carefully, with rigor, and with integrity. Data falsification is inexcusable and is damaging to the investigator and the scientific community as a whole.

CONCLUSIONS

Surveys are designed to obtain information from a target population. Careful planning of specific aims, methodology, data collection, and data analysis allows for useful conclusions that minimize bias and maximize validity. Investigators should consult with survey experts and biostatisticians to fit these guidelines with their target needs. When executed successfully, meticulous survey development and implementation can summarize valuable information about the knowledge, attitudes, and experiences of targeted individuals and groups.

REFERENCES

- Spandorfer PR, Alessandrini EA, Shaw KN, et al. Pediatric emergency medicine research: a critical evaluation. *Pediatr Emerg Care*. 2003;19:293–301.
- Fowler FJ. *Survey Research Methods*. 4th ed. Thousand Oaks: Sage Publications; 2009.
- Asarch A, Chiu A, Kimball AB, et al. Survey research in dermatology: guidelines for success. *Dermatol Clin*. 2009;27:121–131.
- Burns KE, Duffett M, Kho ME, et al. A guide for the design and conduct of self-administered surveys of clinicians. *Can Med Assoc J*. 2008;179:245–252.
- Groves RM, Fowler FJ, Couper MP, et al. *Survey Methodology*. 2nd ed. Hoboken: John Wiley and Sons, Inc.; 2009.
- Levy PS, Lemeshow S. *Sampling of Populations: Methods and Applications*. 4th ed. Hoboken: John Wiley and Sons, Inc.; 2008.
- Fink A. *How to Conduct Surveys*. Thousand Oaks: Sage Publications; 2008.
- Carifio J, Perla R. Resolving the 50-year debate around using and misusing Likert scales. *Med Educ*. 2008;42:1150–1152.
- Stone DH. Design a questionnaire. *Br Med J*. 1993;307:1264–1266.
- Couper MP, Traugott MW, Lamias MJ. Web survey design and administration. *Publ Opin Q*. 2001;65:230–253.
- McFarland SG. Effects of question order on survey responses. *Publ Opin Q*. 1981;45:208–215.
- Willits FK, Ke B. Part-whole question order effects: views of rurality. *Publ Opin Q*. 1995;59:392–403.
- Umbach PD. Web surveys: best practices. *New Directions Inst Res*. 2004;2004(121):23–38.
- Orvis AK, Dellavalle RP. Institutional review board approval for surveys: why it is necessary. *J Am Acad Dermatol*. 2008;59:718–719.
- Singer E, Bossarte RM. Incentives for survey participation when are they “coercive”? *Am J Prev Med*. 2006;31:411–418.
- Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One*. 2009;4:e5738.

APPENDIX. EXAMPLES OF COMMON ERRORS IN SURVEY ITEM DEVELOPMENT AND EXAMPLES OF IDEAL QUESTIONS

Errors to Avoid	Alternative	Example of Flawed Question	Revised Question
Multiple questions in one	Serial questions	Does your institution have template-based, computerized, discharge instructions for patients after acute closed head trauma? 1. Yes 2. No 3. I don't know	Does your institution have template-based discharge instructions for patients after acute closed head trauma? 1. Yes 2. No 3. I don't know Does your institution have computerized discharge instructions for patients after acute closed head trauma? 1. Yes 2. No
Loaded, leading or emotionally charged questions	Questions that are impartial	A recent study of more than 42,000 children less than 18 years old with head trauma recommended that children with a GCS of 14 should have a head CT scan performed after head injury. Given this finding, if a 13-year-old male patient presented after an assault with a brick, and the only pertinent finding on history or examination was a GCS of 14, would you automatically obtain a head CT scan? 1. Yes 2. No	A 13-year-old male patient presents to your emergency department immediately after a closed head injury from an assault with a brick. There was no loss of consciousness, no vomiting, and he is currently asymptomatic. The only finding on examination is a GCS of 14. Would you automatically obtain a head CT scan? 1. Yes 2. No
Questions with assumptions	Questions without assumptions	In the past year, have you admitted at least 1 patient younger than 2 years to your pediatric intensive care unit after a closed head injury? 1. Yes 2. No	Part 1. Does your hospital have an intensive care unit that accommodates patients younger than 2 years? 1. Yes 2. No Part 2. In the past year, have you admitted at least 1 patient younger than 2 years to an intensive care unit at your hospital after a closed head injury? 1. Yes 2. No
Questions with partial answers	Questions with complete answers	What best describes your work setting? 1. Academic pediatric emergency department 2. Community emergency department	What best describes your work setting? 1. Academic hospital general emergency department 2. Community hospital general emergency department 3. Pediatric emergency department within a children's hospital 4. Pediatric emergency department within a general hospital 5. Other (please specify) _____
Trademarked names	Generic names (unless trademarked name is more familiar, in which case include both polyglactin suture (Vicryl)	When a patient has a headache after a minor acute closed head injury, which of the following oral analgesics have you used? (select all that apply) 1. Tylenol® 2. Motrin® 3. OxyContin® 4. Other (please specify) 5. I have never used oral analgesics for this purpose	When a patient has a headache after a minor acute closed head injury, which of the following oral analgesics have you used? (select all that apply) 1. Acetaminophen 2. Ibuprofen 3. Oxycodone 4. Other (please specify) 5. I have never used oral analgesics for this purpose

(continued on next page)

APPENDIX. (continued)

Errors to Avoid	Alternative	Example of Flawed Question	Revised Question
Acronyms and colloquialisms*	Full sentences and proper terms (with acronyms when appropriate)	If a child <18 years old has LOC after a minor closed head injury, I always scan their head. 1. True 2. False	If a child <18 years old has LOC after a minor closed head injury, I always perform a CT scan of the head. 1. True 2. False
Using ambiguous terms	Clearly defining a term for the purpose of the question	I regularly see pediatric patients <18 years old with minor closed head injury 1. Yes 2. No	In the past year, I have cared for 12 or more patients <18 years old with minor closed head injury. 1. Yes 2. No
Choices that use percent	Either use words or a combination of words and percent	How often do you obtain skull radiographs in children younger than 2 years with minor closed head injury and a GCS of 15? 1. Never 2. 1% to <26% 3. 26% to <51% 4. 51% to <76% 5. >76%	How often do you obtain skull radiographs in children younger than 2 years with minor closed head injury and a GCS of 15? 1. Never 2. Rarely (1% to <34%) 3. Sometimes (34% to <66%) 4. Often (67% to <100%) 5. Always (100%)
Questions that deal with social desirability	Questions that are neutral and factual	Have you ever obtained a CT scan of the head that you thought was unnecessary just because a parent or guardian asked you to? 1. Yes 2. No	Emergency department providers have previously endorsed many different factors that are included in the decision to obtain a CT scan of the head after minor closed head injury. Has parental or guardian request for CT scan ever been included in your decision-making process? 1. Yes 2. No

*For the purposes of the examples in this appendix, acronyms have been defined in a footnote. In an actual survey, ensure that acronyms are clearly defined in the text or are explained to the survey participant in advance.

GCS indicates Glasgow Coma Scale; CT, computed tomography; LOC, loss of consciousness.