Threats to measurement validity in self reported data can be overcome

Angela Mickalide

The article by Evans and colleagues in this issue of the journal (p29) is exemplary of the strengths and weaknesses of descriptive data. The authors succeeded in gathering a broad range of information (for example, knowledge, attitudes, and self reported behaviors on several injury risks) from a diverse group of parents (for example, low income, high income, employed, unemployed) using a pilot tested mailed survey based on previously published data collection instruments. The authors failed, however, in attaining only a moderate response rate despite two attempts at follow up (for example, 67% and 58% from most and less affluent families, respectively). In addition, they measured socioeconomic status indirectly through residence; rather, each respondent should have been asked to provide household income for cross verification. Further, the authors failed to include a neutral option (for example, ‘neither agree nor disagree’) among the forced choice selections (for example, ‘strongly agree’, ‘agree’, ‘disagree’, ‘strongly disagree’), resulting in strikingly reduced response variation. Despite these shortcomings, this research is decidedly representative of the published survey findings of parental knowledge, attitudes, and self reported behaviors concerning childhood injury prevention.

Criticisms of the measurement validity of self reported data on childhood injury prevention practices abound, as aptly articulated in my colleague’s opinion. Validity is defined as the ‘adequacy with which the method of measurement does its job—how well does it measure the characteristic that the investigator actually wants to measure?’ The measurement validity of mailed surveys and telephone interviews are often viewed as suspect as they may elicit socially desirable responses resulting in over-reporting of safety behaviors and under-reporting of risky behaviors by parents and caregivers.

Although self report methods admittedly may not be as rigorous as well designed observational surveys in measuring behavior, they have several advantages. Data are much less expensive to amass, more expedient to collect, and provide information otherwise not accessible directly through observation—namely, respondent’s knowledge, attitudes, and opinions. Further, this research method has been used extensively in the injury prevention field, specifically and in the public health field, generally (for example, National Health Interview Survey, Behavioral Risk Factor Surveillance Survey, US Census).

Research findings, in turn, have underpinned the design and implementation of countless safety interventions. To discount the validity of self reported data would be tantamount to positing that injury prevention practitioners have been intervening for years in the safety habits of families on, at best, shaky evidence. I believe that self report surveys can yield valid data, providing the researcher acknowledges possible sources of error, and if certain precautions are taken.

(1) Strive to attain face and consensual validity

Face validity refers to whether the questions posed appear likely to yield the information one wants to measure. For example, the question, ‘In the past year, did you always wear a bicycle helmet while bicycling?’ seems to adequately assess the time frame and frequency of use of this safety device. A survey has poor face validity if there are many ‘don’t know’ responses to such questions. In turn, consensual validity is achieved when a number of experts agree that a measure is valid. To ensure consensual validity, researchers should solicit the advice of other injury prevention experts (for example the National SAFE KIDS Campaign (NSKC) calls upon its Technical Advisory Board) to critique the survey questions.

(2) Be cognizant of the survey’s intrusive nature

Researchers must recognize that questions concerning childhood injury prevention practices are sensitive, because they call into question parental competence, and in some instances, adherence to state or local laws (for example, bike helmet and child safety seat legislation). Most parents want desperately to see themselves as good protectors of the health and wellbeing of their children. If questions along the lines of ‘Do you always supervise your child near water?’ force a ‘yes/no’ answer choice, most parents will respond in the affirmative. Instead, researchers should provide respondents with forced choice, ordinal responses such as ‘always’, ‘almost always’, ‘sometimes’, ‘seldom’, ‘never’ so that parents and caregivers can admit to having occasional supervisory lapses without feeling embarrassed or inadequate. Posing questions in this manner, and contrary to original expectations, the NSKC found that the majority of parents were
(3) Write clear questions
Instead of simply asking ‘Does your home have emergency phone numbers posted near your telephone?’ researchers should provide examples such as ‘like police, fire, and doctor’ to ensure that respondents understand what are considered emergency numbers. In addition, researchers should gain permission to replicate previously validated questions from colleagues’ surveys. To eliminate unclear wording and questions that yield little response variability, the instrument should be pilot tested with a sample of at least 40 parents or caregivers. Ideally, investigators should consider collaborating with a professional survey research firm to write, pilot test, and field the instrument.

(4) Maximize construct validity
To assess ‘the extent to which a particular measure relates to other measures’, compare results to previous research findings, injury patterns, and a theoretical framework. If survey responses approximate those found in other research, then there is a high likelihood that the questions are valid. For example, the NSKC’s Family Safety Check survey revealed that one in three families still did not buckle up on every car ride, a finding consistent with data from the National Highway Traffic Safety Administration. Further, the fact that each year in the US 28 000 children ages 14 and under are treated in hospital emergency rooms for scald related injuries is consistent with the Family Safety Check survey finding that two out of three parents reported that they did not have their home hot water temperature set at a level low enough to prevent scald burns or did not know the temperature. Use of a theoretical framework, such as the Health Belief Model to guide research also can help to assess measurement validity. For example, if a researcher posits that a positive association exists between the perceived threat of residential fire and use of smoke detectors in the home, and the self reported data support this hypothesis, then greater confidence can be placed in the validity of the measures used.

(5) Use trend analysis
Survey research also provides trend analyses that can minimize the likelihood of over or under-reporting by respondents. For example, the NSKC posted several identical questions in its 1987, 1992, and 1995 surveys to measure trends over time (for example, when you think about things that could happen to your children, what sorts of things do you worry about the most?). Childhood injuries (as opposed to drugs and violence) have become an increasingly graver concern in the minds of American parents over the past decade, showing increased societal awareness of the leading cause of pediatric deaths and injuries.

An ideal research design couples self report measures with observational measures. Yet this is not always feasible given shortages in time, money and staff, and the invasiveness of mounting a ‘hidden camera’ in every home, on every street corner. Despite the fact that self report methods are not the most rigorous, the less the will continue to be relied upon in numerous arenas from political polling to market research to health care surveys. Injury researchers should forge ahead in collecting mailed and telephone survey data and refuse to let the quest for perfection be the enemy of good, proactive planning in prevention research.

The author thanks Deborah Glik, ScD, Peter Scheidt, MD, Andrea Glien, ScD and Paul Harvard, PhD.