Recall Bias can be a Threat to Retrospective and Prospective Research Designs

E Hassan

Abstract

Recall bias represents a major threat to the internal validity of studies using self-reported data. It arises with the tendency of subjects to report past events in a manner that is different between the two study groups. This pattern of recall errors can lead to differential misclassification of the related variable among study subjects with a subsequent distortion of measure of association in any direction from the null, depending on the magnitude and direction of the bias. Although recall bias has largely been viewed as a common concern in case-control studies, it also has been documented as an issue in some prospective cohort and randomized controlled trial designs. The aim of this paper is to address recall bias in selective studies employing retrospective and prospective designs and present some key methodological strategies to consider in analytic research using reported data in order to avoid or minimize recall bias.

INTRODUCTION

Bias is defined as deviation of results or inferences from the truth, or processes leading to such deviation[1]. It is the ultimate consequence of introducing systematic errors at any stage of investigation[2]. The term "bias" is sometimes referred to the lack of internal validity which is of central importance in epidemiologic research[3]. Among the several classifications of biases in the literature is the classification by Kleinbaum et al., who classified biases into three main classes: selection bias, information bias, and confounding[4]. Unlike confounding bias, selection and information bias cannot be corrected or controlled for after the completion of a study[1]. Therefore, it is critical during the planning stage of research to address the possible sources of these two biases and consider expedient strategies to avoid or at least minimize them.

Recall bias is a classic form of information bias[1]. It represents a major threat to the internal validity and credibility of studies using self-reported data[5]. According to Sackett's catalog of biases in analytic research, recall bias can be introduced in the data collection stage of investigation[6]. It arises when there is intentional or unintentional differential recall (and thus reporting) of information about the exposure or outcome of an association by subjects in one group compared to the other. This differential recall can lead to differential misclassification of the study subjects with regards to the exposure or outcome variable[1]. Recall bias of sufficient magnitude can depart the estimated measure of effect size either towards or away form the null, depending on the proportions of subjects misclassified. The risk estimate is biased away from the null if more cases incorrectly report being exposed or more exposed individuals incorrectly report developing a disease in case-control and prospective cohort studies respectively[7].

Recall of information depends entirely on memory which can often be imperfect and thereby unreliable[8]. People usually find it difficult to remember or accurately retrieve incidents that happened in the past because memory traces in humans are not but poor versions of the original percept[9]. Research tells us that 20% of critical details of a recognized event are irretrievable after one year from its occurrence and 50% are irretrievable after 5 years[10]. Several mental processes contribute to this characteristic of humans' memory that often threatens the validity of self-reported data in analytic research: some details of an event may go unnoticed by the brain and thus never be stored in memory; memory tends to distort perception in systematic ways; repeated retrieval of already stored events may add new information as facts and thus events are re-stored in the brain in an altered fashion[11]. Given this complex non-dependable process of storing incidents, it has been concluded that the accuracy of recall in humans significantly depends on the

time interval between the event and the time of its assessment: the longer the interval, the higher the probability of incorrect recalls[12].

In general, recall bias can highly be expected in studies using reported data if one or more of the following conditions exist: the disease/event under investigation is significant or critical such as cancer or congenital malformation ; a specific exposure is preconceived by the patient as a risk factor of a high burden disease such as attributing increasing incidence of leukemia in a geographic area to electromagnetic fields produced by a nearby power lines; a scientifically ill-established association is made public by the media such as publicizing the ill-evident linkage between artificial light and risk of breast cancer; or the exposure under investigation is socially undesirable such as reporting of illicit drugs intake [12,13,14].

Although recall bias has largely been viewed as a constant major concern in case-control studies, it has also been documented as an issue in specific conditions of prospective cohort and clinical trial designs. The objectives of this paper are: to address recall bias in retrospective and prospective research designs and present key methodological strategies to consider in the design of research using reported data in order to avoid or minimize recall bias.

RECALL BIAS AND CASE-CONTROL DESIGNS

Participants in case-control studies mainly rely on their memory to identify what in the past might have caused their current disease which is most often of long latency. Because human memory is frequently imprecise, recall bias (According to Grimes and Schulz, 2002)[1] is commonly believed to be "pervasive in case-control studies". The presence of disease is presumed to act as a stimulus that affects both the patient's perception of the causes and his search for possible exposure to a hypothesized risk factor[3]. Therefore, the recall of remote exposures in case-control studies is commonly presumed to be differential among study subjects depending on their disease status[15]. Data, even about irrelevant exposures, are often remembered better by cases or/and underreported by controls[16]. This trend in exposure recall tends to inflate the risk estimate in case-control studies[7] (see Figure 1). Also, recalling the exact timing of exposure which is often important in determining temporality of an association and in estimating induction period of a disease can be differential among exposed cases and exposed controls[17].

Figure 1

Figure 1: Recall bias in case-control studies

Cases recall past exposure better than controls & controls tend to underreport their past exposures: ↑ a or ↓ d in case-control studies → ↑ OR

Logically, if recall of past events is unreliable if reported by subjects in case-control studies, then recall bias is more likely to be greater if information on past exposures is collected from a proxy[18]. This contention is supported by the conclusions of many case-control studies about the unreliability of responses from proxy respondents. For example, the evidence provided by two studies using proxy responses for two different associations: the use of herbicide 2, 4-dichlorophenoxyacetic acid and risk of non-Hodgkin's lymphoma; exposure to hazardous waste and risk of unfavorable respiratory health outcomes, was negated when the cases responded for themselves[19, 20].

Recall bias has often been cited in case-control studies on congenital malformations or cancers in infants[17]. As noted previously, parents of children with serious congenital malformation have the incentive to recall all possible past events that could have caused the disease; whereas parents of healthy children lack such motivation. This is clearly demonstrated in the study by Rockenbauer and associates, 2001[21] which found that reported-data on drug intake during pregnancy by mothers interviewed few months after birth showed evidence of recall bias when compared to drug intake data recorded in a log-book by obstetricians during pregnancy. The sensitivity of exposure reporting was higher for cases than for controls. That means the proportion of truly exposed mothers correctly classified in the study was higher in cases than in controls, indicating better recall by mothers of cases. Furthermore, the noticed lower specificity of self-reported exposure for cases than controls indicates overreporting of the exposure by mothers of cases: the proportion of truly unexposed mothers correctly classified in the study was lower in cases than controls (Table 1). It is interesting to note that the timing of drug intake in this study was reported slightly closer to the time of interview for cases than for controls.

Figure 2

Table 1: Evidence of recall bias: self-reported drug intake data by mothers of infants with congenital malformation given after birth compared to log-book data recorded by obstetricians during pregnancy.

Self-reported drug intake data	Mother of cases		Mothers of controls		Sensitivity	Specificity
	Drug intake data in Log-book Recorded	Not-recorded	Drug intake data in Log-book Recorded	Not-recorded		
Exposed	571	3,095	1,494	649	62% (cases)	86% (cases)
Not exposed	363	18,835	3,856	32,153	28% (controls)	98% (controls)
Total	934	21,930	5,347	32,802		

Adapted from: *Rockenbauer et al. Epidemiology 2001;12(4):461-466.*

On the other hand, another group of investigators studying the same association have reported that recall bias might not be a major concern in case-control studies using parent-reported data as it has often been perceived. This argument received a substantial support from the results of a recent review of empirical studies that assessed the validity of parental reporting in case-control studies on different childhood diseases (leukemia, autistic disease, and sudden infant death syndrome) by using either adequate or gold standard data, such as medical records [22]. The authors asserted in their review that a considerable number of 100 evaluated variables on past exposures suffered from inaccuracies in the reported related information equally by parents of both case and control subjects. Because nondifferential recall errors nearly always tend to depart the odds ratio towards the null value, they cannot account for the positive finding of a research and thus they are insignificant[3]. However, it is important to note that this rule may not hold and a bias away from the null can occur in nondifferential misclassification if the exposure variable has more than two categories[23]. Only a few of the evaluated variables in the review showed evidence of recall bias with a subsequent insignificant differential misclassification. Nevertheless, investigators of case-control studies using parental-reporting are constantly encouraged to consider use a proxy source of reported data if possible to evaluate whether differential reporting by study group has occurred[5,22].

Advocating for the precautionary principle, results from case-control studies in general should be interpreted with caution because the pattern of recall bias frequently encountered in such design tends to inflate the estimated risk attributed to the exposure under investigation and this could potentially yield spurious association.

RECALL BIAS AND PROSPECTIVE COHORT DESIGN

In prospective cohort studies using self-reported data, exposure data are collected before the occurrence of study outcome. Accordingly, prospective cohort design has been largely perceived as an effective strategy to avoid exposure recall bias that is frequently inherited in retrospective designs[24, 25]. However, it has been argued that differential recall of exposure is possible in prospective cohort studies if exposure variable is transient, with short induction period and repeatedly measured over time through self-report: e.g. episodes of anger or stress[26]. In this circumstance, there is opportunity for outcome onset to precede exposure self-report. This phenomenon is more likely to occur if the exposed individual has prior knowledge about the possible outcomes of an exposure. The empirical study by Kip et al[26] addressed recall bias in a prospective cohort study of the association between recurrent ocular herpes simplex virus (HSV) disease and systematic infection and psychological stress as putative risk factors. Findings from this study indicate that self-reported exposure data collected on or after the onset of the disease are more likely to be overreported (recalled better) when compared to the same data collected before the onset of the disease (a standard data collection process in the protocol of any prospective cohort study). This differential reporting can be explained by the concept of rumination bias: people with a disease tend to think harder about their prior exposures than disease free people[6].

RECALL BIAS AND RANDOMIZED CONTROLLED TRIALS

Randomized controlled trials (RCTs) with subjective outcomes may also be contaminated by recall bias if patients enrolled in the trial were not blinded to their treatment allocation. A participants' knowledge about what they receive may influence their reports of related effects, particularly if the outcome data are reported long enough after the fact. The study by Harnack and colleagues[27] provided an excellent example of recall bias in this specific condition of RCT design. The authors examined intervention-related bias in recalling and reporting of food intake in a population of American Indian children enrolled in several elementary schools randomly assigned to a diet intervention program or a control condition. When the authors compared self-reported data of 24-hour dietary intake collected the next day with direct observation of children while eating their school lunches as an objective measure of the outcome, they found that girls in the

intervention schools systematically underreported their dietary intake relative to girls in the control schools. This trend was not found in boys. The authors attributed the differential reporting of food intake by intervention condition to social desirability bias which might be greater among girls in the intervention schools, where healthy eating is emphasized in the classroom curriculum. Recall bias may be most marked in RCTs if people who collect self-reported outcome data are not blinded to treatment allocations[1].

APPROACHES TO MINIMIZE RECALL BIAS

Irrespective of study design, the first step in the process of avoiding any type of bias is the proper definition and articulation of the research question. Consequently, this step will lead to a number of questions that need to be adequately addressed by the investigator during the planning stage of research: what kind of information are required to answer this question in the study in terms of exposure, outcome, and possible confounders; what is the most appropriate method to collect these information; and how to achieve comparable accuracy of data collection between the study groups. According to previous research, the accuracy of recall generally depends on: the degree of required detail bout the exposure or outcome[28]; interviewing techniques and the quality of questionnaire; and to some extent the personal characteristics[7, 12]. All of which are important factors to consider in the planning for recall bias elimination.

IN CASE-CONTROL DESIGNS

Despite the fact that recall bias is a major limitation of case-control studies, a number of methodological strategies documented in the literature can minimize recall bias:

1. Using nested case-control design in which reported data on exposures are collected at baseline and throughout a cohort study, if feasible[29].

2. Choosing newly diagnosed cases because remote diagnosis may lead to reporting of newly adopted behaviors as a consequence of the disease[12].

3. Choosing appropriate control group: Finding the perfect control group in case-control studies can be challenging. Many epidemiologists advocate for using patients with a disease not related to the exposure as valid surrogates for population controls[6, 30, 31]. This suggestion is based on the assumption that diseased controls are similar to the cases in their concern about the possible causes of their disease, thus the comparable accuracy

principle between the two groups is not violated[32]. A limitation of this strategy is the possibility of introducing other type of biases such as sampling bias which may occur when the diseased controls have exposures different from those of the general population[12] . Another limitation is the possibility of choosing controls with a disease that has unknown (unexplored) relationship with the exposure. Another group of investigators advocate for using healthy individuals as controls because they proved to be an adequate reference group in some empirical studies[33]. To avoid this debate, some researchers have suggested the use of two control groups in the same study (if possible): a group of healthy individuals and another of diseased controls. Although the latter suggestion may seem more reassuring, it can give rise to confusion if the results were different between the two control groups[32]. The widely accepted strategy in the scientific community is to choose the most appropriate control group within the study context[32].

4. Using standardized data collection protocols: information about exposure should be collected in the same way and at similar timing for cases and controls[1].

5. Using a well-structured and validated instrument for exposure assessment. The instrument should probe detailed questions about the exposure to help the participants report accurate recalls: the number of exposure events, duration of each event, ect[20].

6. Applying the instrument at similar timing in both study groups[34].

7. Giving the participants enough time before answering to reflect and think through a sequence of events in their life history[10, 35].

8. Blinding the study subjects to the study hypothesis and the specific factors being studied. As an example, questions about exposure of interest can be asked among a long list of questions covering other potential exposures[17].

9. Blinding the data collector/interviewer to the outcome status of subjects and the study hypothesis[1].

10. Verification of exposure reported-data by using a reference criterion (e.g. medical records) or another source of reported-data (e.g. data from a spouse or a twin-sibling)5.

11. Conducting a subgroup analysis by the subject knowledge of the purported association to determine if bias exists in the conducted study36.

IN PROSPECTIVE DESIGNS

1. Using standardized data collection protocols: information about outcome should be collected in the same way and at similar timing for exposed and unexposed.

2. Blinding the participants to study hypothesis and RCTs to treatment allocation

3. Blinding the observer/data collector to the study hypothesis, exposure-status of the participants in cohort or treatment allocation in RCTs.

4. Verification of the self-reported data about the outcome via proxy sources, such as direct observation or use of biological markers.

CONCLUSION

Research including reported data about past experiences will always be threatened by the limitations of the individual's memory and the influence of disease/exposure status on the recalling process in humans. Case-control studies are the most subjected design in analytic research to recall bias. However, differential recall is also possible in prospective cohort studies if exposure status is transient, must be periodically recalled and reported, and ascertainment occurs after symptom onset. Empirical studies suggest that recall bias can be a concern even in randomized controlled trials including subjective outcomes if measurements are collected after a period of time from the incidence of outcomes. To avoid or minimize recall bias while designing similar studies in the future, investigators should consider a number of methodological approaches including: use of standardized well structured questionnaire; blinding subjects and data collectors to the study hypothesis; and using proxy sources of reported data if available.

References

1. Grimes D, Schulz K. Bias and causal association in observational research. Lancet 2002; 359: 248-252.
2. Last JM, 2rd ed. A Dictionary of Epidemiology. New York: Oxford University Press; 1988.
3. Delgado-Rodríguez M, Llorca J. Bias. Journal of Epidemiology and Community Health 2004;58:635-641.
4. Kleinbaum G, Kupper L, Morgenstern H. Epidemiologic research. Belmont, CA: Lifetime Learning Publications, 1982.
5. Basso O, Olsen J, Bisanti L, Karmaus W, The European Study Group on Infertility and Subfecundit, The performance of several indicators in detecting recall bias. Epidemiology 1997;8(3):269
6. Sackett L. Bias in analytic research. Journal of chronic diseases.1979;32:51-63.
7. Ibrahim M, Alexander L, Shy C, Farr S, Deming S. Information bias. ERIC Notebook (Jan/Feb) 2001;13:1-4 Available online at http://hsrd.durham.med.va.gov/eric/notebook.asp
8. Koriat, A. How do we know that we know? The accessibility model of the feeling of knowing. Psychological Review 1993;100(4):609-39.
9. Green M. Eyewitness memory is unreliable. Visual Expert Human Resources. Accessed online on (April 13, 2006) from http://www.visualexpert.com/Resources/eyewitnessmemory. html
10. Bradburn N, Rips L, Shevell S. Answering autobiographical questions: The impact of memory and inference on surveys. Science, New Series 1987; 236(4798):157-161.
11. Fricker D, Reardon E, Spektor D, Cotton S, Hawes-Dawson J, Pace J, Hosek S. Pesticide use during the Gulf war: a survey of Gulf war veterans. National Defense Research Institute; 2000. accessed online on (April 4, 2006) from http://www.gulflink.osd.mil/library/randrep/pesticides_surve y/mr1018.12.appd.html
12. Margetts B, Vorster H, Venter C. Evidence-based nutrition: the impact of information and selection bias on the interpretation of individual studies. South African Journal of Clinical Nutrition 2003;16(3):78-87.
13. Wynder EL. Investigator bias and interviewer bias: the problem of systematic error in epidemiology. Journal of Clinical Epidemiology 1994;47:825-7.
14. Raloff J. Does light have a dark side? Nighttime illumination might elevate cancer risk. Science News 1998;154(16):252.
15. Gabbe B, Finch C, Bennell K, Wajswelner H. How valid is a self-reported 12 month sports injury history? British Journal of Sports Medicine 2003;37:545-547.
16. Choi B, Noseworthy A. Classification, direction, and prevention of bias in epidemiologic research. Journal of Occupational Medicine 1992; 34:265-71.
17. Schüz J, Spector L, Ross J. Bias in studies of parental self-reported occupational exposure and childhood cancer. American Journal of Epidemiology 2003;158(7):710-716.
18. Chouinard E, Walter S. Recall bias in case-control studies: an empirical analysis and theoretical framework. Journal of Clinical Epidemiology 1995; 48(2):245-54.
19. Zahm S, Weisenburger D, Babbitt P, Saal R, Vaught J, Cantor K, Blair A. A case-control study of non-Hodgkin's lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in Eastern Nebraska. Epidemiology 1990;1(5): 349-56.
20. Mohan A, Degnan D, Feigley C, Shy C, Hornung C, Mustafa T, Macera C. Comparison of respiratory symptoms among community residents near waste disposal incinerators. International Journal of Environmental Health Research 2000; 10:63-75.
21. Rockenbauer M, Olsen J, Czeizel A, Pedersen L,

Sorensen H, EuroMAP Group. Recall bias in a case-control surveillance system on the use of medicine during pregnancy. Epidemiology 2001;12(4):461-6.

22. Infante-Rivard C, Jacques L. Empirical study of parental recall bias. American Journal of Epidemiology 2001;152 (5): 480-6.

23. Dosemeci M, Wacholder S, Lubin J. Does nondifferential misclassification of exposure always bias a true effect toward the null value? American Journal of Epidemiology 1990; 132:746-8.

24. Terry P, Bergkvist L, Holmberg L, Wolk A. Coffee consumption and risk of colorectal cancer in a population based prospective cohort of Swedish women. Gut 2001;49:87-90

25. Flood A, Velie E, Sinha R, Chaterjee N, Lacey J, Schairer C, Schatzkin A. Meat, fat, and their subtypes as risk factors for colorectal cancer in a prospective cohort of women. American Journal of Epidemiology 2003; 158:59-68.

26. Kip K, Cohen F, Cole S, Wilhelmus K, Patrick D, Blair R, Beck R.. Recall bias in a prospective cohort study of acute time-varying exposures: Example from the herpetic eye disease study. Journal of Clinical Epidemiology 2001; 54: 482-487

27. Harnack L, Himes J, Anliker J, Clay T, Gittelsohn J, Jobe J, Ring K, Snyder P, Thompson J, Weber J. Intervention-related Bias in Reporting of Food Intake by Fifth-Grade Children Participating in an Obesity Prevention Study. American Journal of Epidemiology 2004;160(11):1117-21.

28. Loftus E, Smith K, Klinger M, Fiedler J. Memory and Mismemory for Health Events: (In J. M. Tanur; ed. Questions about questions: Inquiries into the cognitive bases of surveys. New York: Russell Sage Foundation; 1991: 102-37.

29. Friedenreich C, Howe C, Miller A. Recall bias in the association of micronutrient intake and breast cancer. Journal of Clinical Epidemiology 1993 ;46 :1009 -17

30. Werler M, Pober B, Nelson K, Holmes L. Reporting accuracy among mothers of malformed and non-malformed infants. American Journal of Epidemiology 1989;129:415-21.[abstract]

31. Kaerlev L, Lynge L, Sabroe S, Oslen J. Colon cancer controls versus population controls in case-control studies of occupational risk factors. Biomedical Central Cancer 2004, 4:15.

32. Wacholder S, Silverman D, McLaughlin J, Mandel J. Selection of controls in case-control studies: types of controls. American Journal of Epidemiology 1992;135: 1019-28.

33. Delgado-Rodriguez M, GdMez-Olmedo M, Bueno-Cavanillas A, Garcfa-Martin M, Galvez-Vargas R. Recall bias in a case-control study of low birth weight. Journal of Clinical Epidemiology 1995;48 (9):133-40.

34. Horvath FW. Forgotten unemployment: recall bias in retrospective data. Monthly Labour Review, Research Summaries (March)1982: 40-44.

35. Auriat N. My wife knows best": A comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium Population Register. Public Opinion Quarterly 1993; 57:165.

36. Wartenberg D. Problems in conducting epidemiological studies. FACSNET; 1993. accessed online (on April 19, 2006) from http://www.facsnet.org/tools/ref_tutor/epidem/problems.php3

Author Information

Eman Hassan, MBBCh, MHSc, PhD (Candidate)

Department of Health Care and Epidemiology, Faculty of Medicine, University of British Columbia

Eman Hassan, MBBCh, MHSc, PhD (Candidate)