Invited Commentary

Invited Commentary: Recall Bias in Melanoma—Much Ado About Almost Nothing?

Olaf Gefeller

Initially submitted August 14, 2008; accepted for publication August 28, 2008.

Recall bias has been given considerable attention in textbooks and methodological research because of its potential to jeopardize the validity of epidemiologic results. Case-control studies on self-reported ultraviolet radiation exposure as a risk factor for melanoma have been described as especially prone to the deleterious effect of recall bias because of the growing public awareness about these risks. Using an ideal test-retest design in a large nested case-control study, Parr et al. (Am J Epidemiol. 2009;169(3):257–266) examined to what extent recall bias in melanoma risk factors is actually identifiable and which consequences its presence has on effect estimates of these risk factors. They found only minor indications of recall bias, showing an inconsistent overall pattern and a quite negligible effect on risk estimates. Recall bias was not observed in those exposures where it was most expected (solarium use and other ultraviolet radiation-related exposures). Their findings cannot be used as an argument that future case-control studies in melanoma epidemiology should be avoided because of the biasing effect of recall bias.

bias (epidemiology); dermatology; epidemiologic methods; melanoma; ultraviolet rays

Abbreviation: NOWAC, Norwegian Women and Cancer Study.

The evaluation of potential bias is a key topic in epidemiology. All observational research is prone to lots of sources of bias that can seriously distort the truth. Epidemiologic textbooks and courses devote tremendous efforts to sensitize the novice about the problem of bias. Typically, at the end of a classic epidemiologic education, there are 2 types of students: One does not believe any empirical finding from an observational study and is afraid that one of the dozens of sources of potential bias has produced spurious associations or has masked a true exposure-disease-relation; the other type of student gets quickly rid of the alleged ballast of theoretical discussions about bias and focuses on analyzing epidemiologic data, hopefully implementing techniques to control for bias whenever possible.

RECALL BIAS IN GENERAL

One specific form of bias that has been given considerable attention in nearly all traditional and modern epidemiologic textbooks, however, does not belong to those forms of bias that can be easily taken into account during analysis. Its presence in retrospective research was recognized quite early. A classic example is given in Social Origins of Depression: A Study of Psychiatric Disorders in Women by Brown and Harris (1), who describe a study by Stott in 1958 (2). Stott compared data from interviews with mothers of children with Down syndrome and mothers of normal children. He found that, in the first group, significantly more shocks during pregnancy had been reported by the mothers and concluded that socioemotional factors played an etiologic role in the disease process. Later research, of course, clarified that chromosomal abnormalities were the cause, and mothers of children with Down syndrome probably redefined ordinary events as shocks in an effort to explain what had happened.

Such a differential memory of exposure history in the study subgroups to be compared has often been termed “recall bias.” It is important to note that recall bias is not
RECALL BIAS IN MELANOMA

equivalent to memory failure itself. If an incorrect memory of past exposure, which is inevitable to some degree in all retrospective studies, affects the groups to be compared to an identical extent, recall bias will not occur. This form of bias is a special variant of an information bias that can be thought of as a form of differential misclassification. It can arise in all studies using retrospective exposure ascertainment when study subjects are aware of their disease status and show a (intentional or unintentional) tendency to report past exposure differentially depending on their disease status. Both downplaying and exaggeration of individual exposure have been offered as the typical reaction pattern of diseased study subjects when recollecting their exposure history. In epidemiologic practice, such a situation may occur in many case-control studies.

Several methodological papers (3–7) have addressed this form of bias and analyzed its consequences, but no clear-cut message could be drawn from these analyses. The direction and magnitude of the bias depend in a complex way on quite a number of parameters of the specific study framework. Recall bias, like differential misclassification in general, is, therefore, the nasty type of bias in epidemiology: Everyone knows that it might be present, but usually nobody knows whether this is actually the case and what consequences its presence will have.

Despite the attention given to recall bias in epidemiologic textbooks and methodological papers, empirical research focusing on recall bias itself is quite limited. Only scattered reports trying to assess the presence, direction, and magnitude of recall bias in specific settings can be found in the literature. There is much more speculation in a lot of modern case-control studies about potential recall bias than sound data corroborating its existence and delineating its effect on the study results. The reason for this discrepancy is simple: It is quite difficult to study recall bias in practice. Strictly speaking, the only methodologically appropriate approach consists of a test-retest design in cases and controls where the first assessment had been made prior to the diagnosis of the cases and the second assessment after the diagnosis (the controls have to be studied to disentangle recall bias from nondifferential misclassification that can also show some temporal trend in exposure reporting).

RECALL BIAS IN MELANOMA

Not all areas of epidemiologic research are considered to be equally prone to the problem of recall bias. When exposure information has to be retrospectively ascertained over a wide-stretching time period and can be based solely on self-reports of study subjects without the opportunity of external verification, recall bias is regarded as a serious threat to the validity of the corresponding studies. Such concern has especially been expressed for those exposures that attract public attention as the knowledge of a harmful effect of an exposure is a prerequisite for study subjects to change their reporting behavior of individual exposure. In melanoma epidemiology, examining those risk factors related to ultraviolet radiation exposure in case-control studies is thus notably problematic with respect to the presence of recall bias, since all the conditions given above hold for these factors.

The established effect of ultraviolet radiation exposure on the occurrence of melanoma (8) has received intensive media coverage during the last 2 decades. Gandini et al. (9) observed in their meta-analysis on the effect of total sun exposure significant heterogeneity between effect estimates published before and after 1990 and attributed this finding to a growing impact of recall bias. They also explained discrepant results from population-based and hospital-based case-control studies by the presence of recall bias when hospital-based control groups were interviewed. They argued that controls with dermatologic diseases—constituting the majority in published hospital-based case-control studies of melanoma—may be more aware of the general effect of ultraviolet radiation than population controls are (9).

Public awareness resulting in recall bias has also been made responsible for the lack of an effect of ultraviolet radiation exposure and solarium use on melanoma risk in a recent large European case-control study (10). In a separately published additional analysis, evidence has been presented that recall bias is a plausible explanation for the phenomena observed in that study (11).

Contrary to the many investigations evaluating reproducibility of information in melanoma risk factors (12–21), only a few attempts to study recall bias on melanoma empirically can be found in the literature (22–27). These findings have been briefly summarized in the work by Pfahlberg and Gefeller (28) and in the Discussion section of Parr et al. (29) in this issue of the Journal. All of these previous investigations suffer from limitations: very limited sample size (25, 26), slight changes in the wording of questions in the first and second exposure ascertainments (23, 27), or indirect approaches not using a test-retest design (22, 24).

RECALL BIAS IN MELANOMA BASED ON THE NORWEGIAN WOMEN AND CANCER STUDY

This unsatisfactory situation is changed by the publication of the sound work by Parr et al. (29). Based on the Norwegian Women and Cancer Study (NOWAC) cohort, a nested case-control study of 162 melanoma cases and 1,242 matched controls was constructed to analyze recall bias in an ideal setting. Comprehensive exposure data had been ascertained at enrollment and were regathered several years later in 2004. Then, Parr et al. compared shifts in responses between melanoma cases and controls, finding little evidence of a strong differential recall of exposure history. They observed some changes in reported exposure to melanoma risk factors between the first and second assessments, but these changes occurred in cases and controls. A systematic shift in cases only was confined to the variable “skin color after chronic sun exposure,” a finding that is hard to interpret. In addition, a significantly higher number of nevi in the second assessment was reported by cases and controls with the magnitude of shift being larger for cases than for controls. For all other risk factors, especially for the interesting ultraviolet radiation-related exposures, no evidence of recall bias could be found. Parr et al. conclude that
“indications of recall bias were found . . . but . . . the results were inconsistent for the different exposures” (29, p. 257), which correctly summarizes the surprising pattern of results. Recall bias was not found in those exposures where it was most expected (solarium use and other ultraviolet radiation-related exposures) and, where it was found, it cannot be easily interpreted.

In addition to presenting these data on changes in exposure recall, Parr et al. (29) also analyzed whether the use of exposure information from the first (prospective) assessment leads to different risk estimates than the use from the second (retrospective) exposure assessment. This additional analysis addresses the magnitude of the effect of recall bias on study findings when using such imperfect exposure data. Overall, the differences between the results were quite negligible. The same risk factors were identified in both approaches. Although the estimates of different risk categories showed some fluctuations, no consistent pattern of the discrepancies could be identified. A limitation of this analysis was, however, that the risk estimates in some strata of the exposure variables were quite imprecise because of the low number of cases in these categories. Nevertheless, even an indication that the use of retrospectively gathered exposure data will produce different conclusions on melanoma risk factors was not observed.

WHAT SHOULD BE DONE IN THE FUTURE?

NOWAC offers an ideal framework for studying melanoma risk factors in a prospective manner. The first analysis published in 2003 (30) suffered, however, from a low number of only 187 melanoma cases, limiting a more detailed evaluation of risk profiles. A second analysis after an extended follow-up period has been announced (31). A continuation of efforts to use the NOWAC framework for future studies is important for melanoma epidemiology and therefore strongly encouraged. This cohort study approach has the great advantage that nobody will criticize the findings using the argument that recall bias may have influenced the results. A repetition of the current methodological substudy on recall bias comprising a larger case group to raise the statistical power is also highly welcome to sort out myths from facts in this area.

The case-control approach has, however, been the dominating design in melanoma epidemiology for decades. Hundreds of such studies have assembled our current knowledge about the factors influencing melanoma risk. Despite all methodological shortcomings of the approach, the results of these investigations have guided us well to derive strategies for primary prevention of melanoma. Recently, concerns have been raised about whether case-control studies can still be used in this area because of potential recall bias resulting from the growing public awareness about melanoma risk factors (11). I am not sure whether the findings of the present study in this issue of the Journal will end this epidemiologic discussion. Recall bias is the nightmare of all epidemiologists conducting case-control studies and will remain a hot topic of controversial discussion in the melanoma community. However, some nightmares suddenly lose much of their scare if one awakes in the morning and realizes that all was just a bad dream. In some way, the data by Parr et al. (29) offer such a cozy morning feeling, as they show that the extent of recall bias is much lower than generally expected and that the consequences for the analysis of melanoma risk factors are minor. Definitely, their findings cannot be used as an argument that future case-control studies in melanoma epidemiology should be avoided because of the biasing effect of recall bias.

ACKNOWLEDGMENTS

Author affiliation: Department of Medical Informatics, Biometry, and Epidemiology, University of Erlangen-Nuremberg, Erlangen, Germany.

This work was supported in part by a grant from the German Research Foundation (DFG Ge 637/4-4).

Conflict of interest: none declared.

REFERENCES