Reviewer's report

Title: Concordance of self-reported and medical chart information on cancer diagnosis and treatment

Version: 2 Date: 10 December 2010

Reviewer: Kelly-Anne Phillips

Reviewer's report:

The study by Gupta et al reported in this manuscript describes the concordance between self-reported and medical chart information in a large-population-based cohort of Chinese women with breast cancer. Previous studies of this issue have shown that self-reported data on initial breast cancer treatment are generally reasonably reliable when compared with data abstracted from medical records. Most previous studies have taken place in western health systems, and so whether the findings can be extrapolated to other health care settings is uncertain. This study is a potentially important addition to the existing literature because it is set in a different cultural and health care setting.

Major Compulsory Revisions

The manuscript currently lacks a discussion of how this study compares and adds to others that are already published. The authors should include a discussion of the literature specific to this topic, including, but not necessarily limited to the published studies that have validated self-reported breast cancer treatment data by Phillips J Clin Oncol 2005, Maunsell J Clin Epidemiol 2005, Schootman J Clin Epidemiol 2005, Lui Breast Cancer Res Treat 2010.

The authors should clarify whether the “retired medical professionals” who conducted the interviews were the same individuals who abstracted the medical record data and if so which occurred first. In order for the validity of the self-report data to be confirmed, the individuals who interviewed study participants should not have had access to the data abstracted from the medical records at the time of interview.

The authors state that the time-gap between the in-person interview and the date of diagnosis was a median of 189 days but that the time period of medical chart information abstracted was a median of 14 days after diagnosis. The authors should define the term “date of diagnosis”. Was this the date of the initial biopsy that revealed malignancy or the date of definitive breast surgery or something else?

The fact that the medical record data was only abstracted for the first 2 weeks after “diagnosis” is problematic. Presumably chemotherapy and endocrine therapy were usually not delivered within the first 2 weeks after diagnosis. Thus, is the reader to assume that the individuals who abstracted data from the medical records considered a documented intention to deliver therapy evidence that that
therapy was subsequently delivered? If so this should be clearly stated in the methods section. Presumably in at least some cases, the actual treatment given, and that which patients recalled at 6 months, differed from the initial treatment plan? In such cases the validity of self-report would tend to be underestimated. In my view this is a major limitation of the study, although the authors have acknowledged this.

Also, this study collected self-report data at a very early timepoint after diagnosis (ie 6 months). Presumably many patients at this time would still be undergoing adjuvant treatments, which would tend to aid recall. This should be discussed.

The study did not attempt to validate stage at diagnosis despite the fact that this is a critical factor necessary for epidemiological and quality of care studies. This should be discussed.

In the paragraph on “Concordance in reporting the occurrence of disease-related events”, the authors state that “patients were able to report the dates of receiving these treatments within at least 3 days of the date recorded by the medical chart. “ However the methods section (“statistical analysis”) and table 3 seem to indicate that dates were only coded by month and year, so how can the authors know that the reports were accurate to within days?

Minor Essential Revisions

The radiotherapy recall data described in the first paragraph of “Concordance in reporting the occurrence of disease-related events” should be included in table 3.

In the first paragraph of the discussion, the sentences “the patient questionnaire invariable described.......patient survey relative to the medical chart” describes results that have not appeared previously in the manuscript. This information should appear in the results section. Also the logistic regression that was presumably done to determine whether demographic factors affected concordance should be described in the statistical methods section of the manuscript.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests