Author's response to reviews

Title: Consistency and precision of cancer reporting in a multi-wave national panel survey.

Authors:

Anna Zajacova (zajacova@uwyo.edu)
Jennifer B Dowd (jennbeamdowd@gmail.com)
Robert F. Schoeni (bschoeni@umich.edu)
Robert B. Wallace (robert-wallace@uiowa.edu)

Version: 2 Date: 30 May 2010

Author's response to reviews: see over
We thank the two reviewers for their careful reading of our manuscript and for providing useful comments. Below are the comments, copied verbatim from the review, and our responses.

**Referee #1, Dr. Curado**

1. **When more than one site is reported by the interviewers they used the “primary site.” This is unclear once if the patient has a multiple primary tumour (synchronous or asynchronous) it should be mentioned as first or second tumour because it has an impact increasing patient awareness about cancer.**

   We revised the text to explain the reporting of cancer type better. In the first paragraph of page 9 we now write "In 2005, the survey asked for the first time about the type of tumour: “What type of cancer (do/did) you have? In what part of your body [is/was] it?” The cancer sites included breast, colon, lung, lymphoma or leukaemia, melanoma, prostate, skin (not further specified), uterine, ovarian, cervical, and other. Respondents could report one or more cancer sites; however, over 99.5% of cancer survivors reported only one type of cancer. For the 20 individuals who mentioned a second cancer site, we use the first-mentioned type."

2. **The manuscript addresses the consistency and accuracy on self reporting cancer diagnosis through a self report interview. It is a valid methodology that can be used to access patients knowledge on cancer diagnosis and validate consistency on cancer. The level of inconsistency were high 30% but it is unclear why it was so high once this population is being followed since 1968 is this is an cohort effect?**

   Our analyses cannot determine the causes of the observed inconsistencies. However, the PSID is representative of the US population and therefore the results should be generalizable to similar national surveys. Specifically, the PSID sample and data are of very high quality; the generalizability of the PSID results has been repeatedly upheld in multiple subject domains. The PSID health data in particular have been found comparable to those of the leading cross-sectional US survey, the National Health Interview Survey (1). In addition, analyses of attrition in PSID imply that the gradual changes in the sample due to loss to follow-up have not affected the representativeness of the sample (2). Re-estimating the analyses with weights that correct for most departures from representativeness have led to the same substantive conclusions as unweighted results.

   One reason that prompted this question may have been that we did not adequately describe the study design in which the cohort has been replenished by children and grandchildren of the original sample members, so our results would not reflect a cohort effect.

   We have paraphrased this reply in the "Data" section, adding text to better describe the sample.
3. The limitations were addressed. The interview is time consuming and it is a dependent factor to reduce accuracy. I wonder how many patients asked to stop the interview due the time or how many has to be rescheduled.

Very few patients ask to stop the interview. The wave-to-wave response rate has been over 94% each wave. PSID is a household survey so if one household member chooses not to cooperate, another member may complete the interview, increasing the likelihood that the unit remains in the panel. While the interviews are long, they are not an impediment to collecting information - in contrast, longer interview at a previous wave in PSID are actually associated with a higher likelihood of remaining in the panel (3).

In addition, item non-response for cancer has been less than 1% (1). This high compliance with the extensive interview likely results from the long involvement of the respondents with the PSID and their understanding of the value of this survey.

We added some information about these issues in the Methods section.

4. Self reporting studies are instruments with limited value in cancer for the moment perhaps a special strategy should be used to evaluate cancer. Patients beliefs on cancer diagnosis is not very well understood and this can be a constrain to make public their diagnosis.

Self-reports of cancer have limitations, as described in our manuscript. However, self-reports of cancer are the only cost effective way to collect diagnosis within large national representative population surveys. Moreover, collecting cancer diagnosis in these surveys that also collect detailed information on employment, income, wealth, marriage, child development, and a variety of other social and economic outcomes allows the assessment of the impact of cancer on many aspects of life that cannot be examined with cancer registry data. Therefore, it is important to understand the limitations of the self-reported data such as those collected in the PSID. Future work should attempt to link survey data to registry data to examine these issues more fully.

We paraphrased this response in the "Conclusions."

In addition, we added a comment about the respondents' potential mistakes in understanding their diagnosis and the consequences of these mistakes in the discussion section. In particular, we expanded a paragraph on pages 14-15:

"The consistency of cancer reports may be influenced by two opposing tendencies. On one hand, cancer is a serious, life-threatening disease. The severity and salience of the illness may increase the likelihood that respondents will report cancer with high consistency and precision (4-5). On the other hand, cancer remains associated with stigma and discrimination (6-7), and respondents may be unwilling to report cancer during an interview, leading to underreporting or inconsistent reporting (8). These competing tendencies might contribute to the inconsistent reports of cancer survivors over time.

Alternatively, it may also be that respondents change their cancer report from positive to negative if they consider themselves cured or because recall issues increase with time since diagnosis. This pattern would lead to underreporting of cancer. In contrast, respondents may
initially misunderstand a positive screening test or a diagnosis of precancerous lesion as evidence of cancer, and later realize that these medical findings does not constitute a cancer diagnosis (9). Overall, little is known about the patients’ knowledge and understanding of their diagnosis, as well as about their willingness or ability to report it accurately.”

Referee #2, Dr. de Sanjose

1. I am surprised by the prevalence differences that there is between the panel survey and SEER data. Registry data, for example from CIFC/IARC/WHO is supposed to be the best available information because a large proportion of the date is confirmed using pathology reports. Maybe there is an over estimation of cancer when people are interviewed personally. Maybe the higher prevalence from the panel may also be due to reporting of pre-neoplastic lesions? Could you please provide some discussion on the issue?

The prevalence estimates from 1999 and 2001 are similar to the SEER data, albeit the SEER data are from a more recent time period. Given the different mode of data collection in PSID and SEER, we found this similarity encouraging. Another useful comparison of the PSID data is with the National Health Interview Survey (NHIS), the leading source of data on the health of the US population. The NHIS data are self-reported like in the PSID, and thus more comparable in the mode of data collection. The NHIS all-site cancer prevalence for adults appears somewhat higher at every wave compared to PSID prevalence (1). For example, in 2005 the prevalence was 0.069 in the PSID and 0.074 in the NHIS. As a result, both the NHIS and PSID estimate higher prevalence rates than in SEER. As mentioned by the reviewer, it is possible that survey respondents are reporting precancerous lesions as cancer. This tendency could be pronounced for cervical cancer, but other cancer types may be overreported as well, i.e. breast cancer when carcinoma in situ is reported. We now discuss this fact in the manuscript.

Additionally, we discovered a source of discrepancy between our estimates in 2005 and SEER data. The PSID cancer question in 2005 included respondents with non-melanoma skin cancers, which are not included in the SEER estimates. Correcting this error brought the PSID estimates close to the SEER data. The new estimates and the explanation of their calculations are in Table 2 and in the results section.

2. Cervical Cancer can also be affected by lack of confirmation. Sometimes women are reported to have a pre-enoplasic lesion leading to cancer that cannot be confirmed afterwards.

We agree with this comment and have revised a paragraph on pages 15-16 to a comment about this potential explanation in the discussion section:

"The findings differ from previous studies in the importance of cancer site as a key predictor of reliable reporting: while previous studies found primary cancer site as the key determinant of under- or over-reporting (8-12), in our models cancer site was not a strong predictor of either consistency or timing precision over time, with the exception of cervical cancer that tends to be reported relatively inaccurately. This cancer type may be particularly likely to be misreported if respondents are reporting precancerous lesions as cancer, and later correcting their cancer occurrence report. The directions of the effects of cancer sites in general, however, were
consistent with the previous studies, whereby colon and breast cancers were reported with high accuracy."

3. In the methods section, it is noted that the prevalence of cancer excludes skin cancer. Contrary, in the results section, 14% of the cancer reported are referred as NM skin cancer. Could you clarify the discrepancy?

While the 1999-2003 surveys excluded skin cancer, the 2005 survey did not. In 2005 respondents were asked to report all types of cancers, including skin cancers. A follow-up question was then asked about the type of cancer. We have changed the text of the manuscript to make this clear.
References.