Reviewer's report

Title: Epidemiology and Outcomes of Undiagnosed Diabetes in Older Women with Breast Cancer: An Observational Cohort Study Based on SEER-Medicare

Version: 1 Date: 2 November 2012

Reviewer: Lorraine Lipscombe

Reviewer's report:

This large, population-based study using SEER and Medicare data sought to compare the influence of undiagnosed versus diagnosed diabetes in breast cancer patients aged 65+ years on (1) cancer stage, (2) treatment and (3) mortality. The study showed that diabetes which was undiagnosed at the time of breast cancer diagnosis (based on a ‘new diagnosis’ within 3 months before or 3 months after cancer diagnosis) was associated with a significantly higher likelihood of advanced cancer stage and increased all-cause mortality. These associations were attenuated when adjusted for prior health care contact but remained significant. This is a well-written paper that addresses an important question regarding the relationship between diabetes and breast cancer. However, there are some methodologic issues and limitations to the interpretation of the findings that should be considered.

Major Compulsory Revisions

1. For the analysis on cancer stage, did the authors look for a significant correlation coefficient between the two chosen variables to evaluate health system contact (health care visits and preventive care score)? Clinically these variables could be highly correlated (i.e. patients who visit doctors more frequently are also likely to have preventive screening tests done) and if this is proven statistically, one should be removed from the model.

2. In the analyses section and figure 2, the authors describe a complicated relationship between 4 different sets of variables: demographic, socioeconomic and clinical characteristics, previously undiagnosed DM, cancer stage and mortality. From the diagram, the authors do not make it clear that previously undiagnosed diabetes is the main predictor of interest – this should be clarified. Second of all, either cancer stage or “demographic, socioeconomic and clinical characteristics” are depicted as being colliders. A collider is a variable that is associated with both the “confounder” variable and the outcome creating a point in the causal diagram where two areas meet head on (as is the case for cancer stage in figure 2). The issue of dealing with such variables is complex – this issue should be discussed in both the methods and discussion session. When one identifies a collider variable one should not include it in the adjusted analysis because this introduces bias, instead one can stratify on this variable or else leave it out of the model altogether. Since the authors are showing the models with and without these variables, it is important that more time should be spent
discussing this issue. Otherwise, including the multiple models are confusing and I would suggest the authors only report results from what they deem to be the most “correct” model.

Good resources are:

3. The authors need to strengthen their hypothesis that undiagnosed diabetes can have ‘direct’ effects on cancer stage independent of differences in health care behaviour and patterns, given that they were unable to account for metabolic factors in this association. The authors should acknowledge in their limitations the lack of information regarding diabetes severity and glucose control. Is anything known about the “severity” of unknown diabetes – are these patients more likely to have very high glucose and A1c levels or high insulin levels and pre-clinical glucose levels? If this is known, such information would help in the discussion of causative pathways for the association seen between undiagnosed DM and later stage/worse prognosis.

4. The authors argue that by adjusting for ‘confounders’, they have demonstrated evidence of a direct (causal) relationship between undiagnosed diabetes and cancer stage. This assertion needs to be tempered, as the possibility of residual confounding due to unmeasured and unknown factors cannot be excluded. Moreover, the substantial reduction in strength of association when adjusting for measured confounders indicates a lack of robustness of this association to confounding. Unmeasured confounders include diabetes severity measures, body-mass index (which can affect clinical detection of early-stage cancers), diabetes treatment/medications (e.g. metformin, which has been shown to affect cancer prognosis), health behaviours, etc... This paper should be viewed as ‘hypothesis-generating’ – further studies will have to evaluate the extent to which the association between undiagnosed diabetes and cancer prognosis is related to health care vs. biologic factors.

Minor Essential Revisions

1. The authors clearly show a strong effect of prior health contact on the magnitude of association between their exposure and outcomes. I have a few questions related to this variable.

a. Given that health care contact is a necessary precursor to having ‘diagnosed diabetes’ based on the study definition (at least one diabetes claim in the 2 years prior), please clarify how 65 patients with ‘diagnosed diabetes’ had no health contact in the 2 years prior.

b. The majority of patients with ‘undiagnosed diabetes’ had some health contact in the 2 years prior to cancer diagnosis – clearly these visits were for something other than diabetes (as there was no claim for diabetes) – is there information on what these visits were for?
c. Did the authors assess for an interaction between diabetes status and health care contact, to determine its role as an effect modifier in the association?

2. The authors should provide a rationale for limiting the cohort to age 65 years or older (due to availability of Medicare data?). Were there other inclusion/exclusion criteria?

3. It is unclear whether all 93% of eligible Medicare patients > age 65 use Medicare solely for their medical needs, i.e. would all doctor visits and diagnostic claims be captured using Medicare data? For readers who are not familiar with the US databases it would be helpful to clarify this point.

4. The authors state that ‘undiagnosed diabetes’ was based on a ‘new’ claim for diabetes (based on a validated algorithm) within the 6-month peri-diagnostic period. There is no mention of whether this definition has been validated (e.g. based on chart review or survey?). Given that the sensitivity of the diabetes case definition is only 74.4%, it is possible that up to 16% of previously undiagnosed cases were actually pre-existing (or prevalent) diabetes cases that did not have a claim for diabetes in the 24-month look-back period prior to diabetes. As shown by the authors, these patients were less likely to have seen a physician in that period. One suggestion would be to conduct a sensitivity analysis with a longer look-back period to exclude those with more remote prior diabetes claims. The authors should mention the lack of validation and limited specificity of this definition.

5. The authors should explain why they did not include a non-diabetes comparison group.

6. The authors adjusted for a measure of prior preventive services, which includes mammography. However, given that mammography is a strong predictor of early diagnosis and stage, it would be useful to also compare rates of mammograms between the 2 groups and how that factor specifically influences the association between the diabetes groups and outcomes, if possible.

7. The authors examine mortality outcomes, but fail to consider post-cancer variables during the follow-up in their model, such as cancer treatment, surveillance, diabetes-related complications. These are important predictors of survival following cancer, and should be mentioned in the limitations if they were not available.

8. Where are the results from the analysis relating to effect of undiagnosed diabetes on cancer stage? The text refers readers to table 1, but table 1 is the table of baseline characteristics.

9. The authors only briefly mention the outcome of time to chemotherapy or radiation, and provide hazard ratios. However there is no information on proportions who received chemotherapy/radiation between groups.

10. The follow-up period appears short for mortality (2-6 years). Please provide information on number and rates of death by group, and follow-up time (median, etc.).

11. The authors should elaborate on the implications of the finding that undiagnosed diabetes is only statistically associated with all-cause mortality, and
not cancer-specific mortality after adjustment for stage and health care patterns. This suggests that patients with undiagnosed diabetes are overall sicker and dying of causes other than their cancer. How are these findings specific to breast cancer? This should be addressed more specifically in the discussion.

Discretionary Revisions

1. The authors conducted a thorough literature review that is well described in the Background. However, there is a lot of detail regarding previous studies that might be better placed in the Discussion section, when comparing this paper with prior research.
2. Figure 3 – it would be helpful to have HR and 95% CI within the graph.

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.