Reviewer’s report

Title: Automated inter-rater reliability assessment and electronic data collection in a multi-center breast cancer study

Version: Date: 2 April 2007

Reviewer: Ronald Lagoe

Reviewer’s report:

General

This manuscript concerns a topic which is greatly in need of additional research, the implementation of electronic data systems in health care. The authors correctly point out that literature concerning electronic data collection is sparse.

The manuscript addresses a limited, but important area of electronic health care data, the actual collection of this information by health care providers. With respect to this subject, the paper focuses on comparison of paper data collection and electronic data collection with respect to a specific area of care, breast cancer treatment. It addresses this subject in a network of six providers within the United States.

The subject selected for this study and the settings for data collection are excellent. Addressing the critical need for additional research concerning electronic data in health care will require the implementation of specific studies concerning well defined topics. This manuscript has the potential to become one of those studies.

-------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The manuscript contains a well constructed methodology and research plan concerning the process of electronic data collection. The process component of the research plan is well described. In this reviewer’s opinion, two subjects require further development.

The first is the need for clarification of the differences between the control samples and the intervention samples of the study. The authors clearly describe data collection for the electronic sample and the methods of evaluating reliability between the two samples. In the opinion of this reviewer, there needs to be more specific description of the control, paper collection sample and how it differs from the electronic sample. This could include additional material describing the two samples at each of the provider sites.

The second subject is the need for quantification of resources required for the paper and electronic samples. In the Background and Table 2, the authors use terms such as “cost effective”, “resources”, and “efficiency” in discussing the strengths of electronic data collection. While the focus of this study was on process and reliability, it seems logical for the authors to generate estimates of the resources required by the two data collection methods. This need not entail the use of exhaustive accounting methods. It could be carried out through budgeting the costs of paper data collection in medical record departments in one or more of the sites and budgeting the costs of electronic data collection used in the study. This exercise would help complete the process of addressing this subject.

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Page 6, Paragraph 3, Line 1, Add an “s” to amount

-------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

As previously noted, this reviewer believes that the process component of the research methodology was well done. At the same time, this evaluation resulted in only a single statistical presentation, Table 1. The
authors might consider whether the evaluation generated material for additional tables involving comparisons of the control and intervention samples and the six sites.

The authors have correctly pointed out that only a limited amount of research is available concerning electronic data collection. They have not indicated, however, whether any of the evaluation tools or other components of that research were used in this study. Such a brief discussion could provide some connection with other published studies.

An additional thought concerns the relationship between the Methods and Results sections. Because this study largely focuses on a process, it is logical that reference to methods not be limited to that section. At the same time, the authors might consider moving some of the discussion concerning the process of evaluation from the Results to the Methods section and allowing the Results section to focus only on the outcomes of the study.

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of importance in its field

**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