Reviewer's report

Title: Availability and Quality of Paraffin Blocks Identified by the Shared Pathology Informatics Network (SPIN): A Multi-institutional Study

Version: 2 Date: 13 March 2006

Reviewer: Matthew R Cooperberg

Reviewer's report:

General
SPIN is a NCI initiative to facilitate sharing of pathology tissue resources among research institutions, under which available resources can be queried remotely. A pilot project is in development among four institutions. This paper reports an important validation step, i.e., confirming whether identified pathology reports can in fact be linked to available tissue slides and paraffin blocks.

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

1. Throughout the introduction and discussion, the authors seem somewhat over-enthusiastic about SPIN, and unfairly derogatory regarding existing institutional tissue banks. As described, SPIN will facilitate identification of tissue resources stored at multiple institutions. It will coordinate rather than replace institutional repositories. The characterizations in the first paragraph, for example, that the utility of existing tissue banks is "limited to a very narrow focus" and that they are "usually abandoned after their funding period terminates" are simply not true; individual institutions are developing tissue banks to answer a wide range of scientific and clinical questions. These have a wide focus and are often funded from multiple sources over the long term; indeed, without existing tissue banks, SPIN would have no raison d'etre. SPIN is not a mechanism for sharing tissue resources per se, and identification of resources at a given institution is by no means a guarantee that that institution will share those resources.

2. There is wide variation in the site-specific methods for query, collection, review, etc.. What is the rationale for not standardizing these methods? The point is made in the conclusions that a future study will use SPIN-only queries, but the wide variation in this study, especially the fact that two institutions used local system queries and two queried SPIN indirectly, requires both justification in the methods and exploration of differences in the results/discussion.

3. Why were methods of specimen collection and review not standardized? For example, outside referrals of rare cases were included for Indiana but not for UCLA. This seems like a source of significant potential bias.

4. One crucial component of institutional tissue banks is correlated clinical data. A list of 100 prostate cancer cases, for example, is of limited utility without details regarding stage, grade, treatment, and ideally outcomes. Is there information accessioned for SPIN outside of what is available in the pathology report? What fields are in the SPIN Excel template?

5. The manuscript in general, and the methods section in particular, is too verbose. I am not sure how the fine details of the query mechanisms, for example, are relevant to the reader.
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

p4. It seems somewhat early to propose that SPIN will coordinate “millions" of tissue samples. There are only four institutions included in this study; even were the project expanded significantly, it would presumably only incorporate academic centers, and perhaps major private sector pathology reference labs. Will the expansion be funded wholly by the NCI? What are the incentives for participation?

p4. Figures are given regarding national pathology resources. Are there any data on the proportion of these specimens which include tumors?

p6. What is the reference / basis for the statement that 10-20% of pathology reports include a cancer diagnosis?

p6. The discussion states that some tumors types initially investigated as rare tumors were discarded due to heterogeneity among sites. This is an important issue to discuss, and should be mentioned early, in methods and/or results not just in discussion. How were the original and replacement rare tumors selected – at random, by concensus of the investigators, etc.?

p6. What does the “n” refer to in “One hundred random n cases”?

p28. Figures 2, 4, and 6 are tables, not figures.

p27. The mix of 2D and 3D bar graphs is distracting and unnecessary. 2D should be used consistently unless there is a specific reason for 3D.

p29. Figure 3 should be formatted as a standard stacked column chart, not a 100% percent chart.

p31. In figure 5, why are there only Harvard cases after 1999?

Discretionary Revisions (which the author can choose to ignore)

p4. For greater accessibility, consider including the SPIN website in the text rather than as a reference.

p31,33. Consider excluding 2005 as the case numbers are quite low.

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

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

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:
I declare that I have no competing interests