DRAFT

Office of the Director
National Cancer Institute (NCI)

Cancer Bioinformatics Grid (caBIG)/National Biospecimen Network (NBN) Pilot Teleconference: Review of NBN Pilot Workflows and High-Level Use Cases

May 6, 2005
8:00 a.m. – 9:30 a.m. EDT

SUMMARY

List of Participants

Mark Adams	Booz Allen Hamilton (BAH)
Harsh Bal	BAH
Greg Eley	BAH
Paul Fearn	Memorial Sloan-Kettering Cancer Center (MSKCC)
Ian Fore	National Center for Bioinformatics (NCICB) Contractor
Mariana González del Riego	Rose Li and Associates (RL&A)
Andrew Hruszkewycz	Organ Systems Branch, NCI
George Komatsoulis	NCICB
Steve O'Krepky	RL&A
Mark Rubin	Dana-Farber Cancer Institute (DFCI)
Julie Schneider	Office of Technology and Industrial Relations (OTIR), NCI
Sharon Settnek	Science Applications International Corporation (SAIC)
John Speakman	MSKCC
Bruce Trock	Johns Hopkins University (JHU)

Introductions and Review of Agenda

Sharon Settnek first invited teleconference participants to introduce themselves. She then reviewed the goals of the teleconference as follows: (1) To review the first iteration workflow diagrams; (2) to review and further develop the first iteration high-level use cases; (3) to understand use case priorities to assist in defining the project scope and timelines; and (4) to discuss the common data elements (CDEs) associated with the prioritized use cases. For additional meeting details, refer to the agenda (Attachment 1).

Prospective Study Workflow Updates

Sharon Settnek displayed the first-iteration (April 18, 2005) workflow diagrams via the virtual Centra session to facilitate group discussion (Attachment 2). She then presented the Prospective Study workflow. As the group considered the diagram, it recognized the need for additional steps. Modifications made to the original prospective workflow diagram appear in bold below and are reflected diagrammatically in Attachment 3.

1) Once the patient is registered by a site (top row) and the site pathology technician processes the tissue, **the site pathologist submits the *pathology data* to the**

administrative core. This last step specifies that detailed pathology annotation will occur. It is a parallel process to that of clinical data depicted in the top row of the workflow.

2) Branching off from the "Site Registers Segment Integrity to Administrative Core" box (bottom row), **the biomarker site research coordinator registers the receipt of the specimen. This then feeds into the Biomarker Workflow.**
3) Once the site research coordinator collects the specimen (bottom row), **the Biomarker Study research coordinator registers the receipt of the specimen. After this step, the Biomarker Workflow is initiated.**
4) Once the site research coordinator collects the specimen (bottom row), **he/she sends the specimen back to the pathology core.**
5) **The arrow from "Site Submits De-Identified Clinical Annotations to Administrative Core" (top row) to "Specimen Sent Back to Site for Resolution" was created in error and therefore removed.**

Retrospective Study Workflow Updates

Several important modifications were made to the Retrospective Study workflow. The discussion about these changes is summarized below. In addition, the resulting modifications appear in bold and are further illustrated in Attachment 3.

1) Bruce Trock noted that the Inter-SPORE Prostate Biomarker Study (IPBS) protocol specifies that an effort will be made by Study Leads to choose one ***set*** of patients and samples on which biomarker studies can be conducted by three Prostate SPORE sites. To address this point, **the "Study Leads Select patients" box (top row) was modified to "Study Leads Selects Patient Set."**
2) Paul Fearn suggested that the step "Site Sends Clinical Data to Admin Core" (bottom row, Retrospective Workflow, Attachment 2) be completed before the block is sent to the tissue microarray (TMA) Core. It should not appear as a separate track. Mark Rubin added that clinical data is not typically associated with a block and that a study identification number is sufficient. He further remarked that this is done to deliberately blind the Pathology Core to the clinical data in order to help preserve the integrity of the study. Bruce Trock further clarified this issue by stating that only a generic subset of clinical data (e.g., treatment data, clinical follow-up data, and biopsy pathology data) would be available to researchers. Study leads would be blinded in this respect to avoid the selection of patients based on outcomes. Therefore, the data displayed to the TMA Core staff would only inform the user about the types of data available with the block. Specific information associated with the patient would not be displayed. As a result of this discussion, **the step "Site Sends Clinical Data to Admin Core," was moved after the step "Study Lead Requests Clinical Data and Specimens from Sites for Patients."** Paul Fearn also highlighted the importance of these steps with respect to the development of CDEs (i.e., determining what data elements would need to be collected) and suggested that the group take a closer look at this section of the workflow during the CDE discussion.
3) Based on the discussion described above, the group added a series of site-specific steps to the workflow (bottom row). Essentially, once the Study Lead requests clinical data and specimens from the sites, **the site pathology technician selects the serum samples and sends them to the Pathology Core. The Pathology Core technician then registers the**

samples and the research coordinator distributes them to the sites for biomarker evaluation. Finally, the biomarker site research coordinator registers the samples upon receipt. At that point, the Biomarker Workflow is initiated.

4) Ian Fore inquired whether TMAs would be constructed in a single study stage or through an ongoing process where blocks are accumulated until a critical point is reached. If TMAs are to be batched, the process would have to be tracked in the system. Mark Rubin agreed that this was a critical issue because, if the blocks and slides are batched, the TMA Core would need to have all the material before the TMAs are constructed. He further added that part of the TMA Core protocol is to confirm that all blocks/slides are received. Based on this exchange of information, **a binomial decision tree including "Blocks Received" and "Blocks Good?" was added after the TMA Core technician registers the blocks received from a given site (middle row).**
5) George Komatsoulis suggested that a messaging acknowledgment be created within the system to verify when shipments are received. Sharon Settnek replied that such notifications would be addressed during the development of use cases. Mark Rubin added that alternate samples would have to be available if some of the original samples are missing or not viable. Bruce Trock agreed with this statement and suggested that alternate samples be submitted by the different sites to account for any nonviable or missing samples. **A binomial decision tree including "Blocks Received" and "Blocks Good?" also was added to the workflow following the registration of samples by the Pathology Core technician (bottom row).**
6) The need to specify the creation of aliquots and their registration at the biomarker site was identified. **Thus, the "Pathology Core Creates Aliquots of Samples" step was added after "TMA Cut Into Slides" (middle row).** The slides or serum aliquots then are distributed to the different sites involved in the evaluation of biomarkers. **At that point, the biomarker site research coordinator registers the receipt of the aliquots.** It was also suggested that a similar series of steps be added to the prospective workflow.

Prospective and Retrospective Biomarker Study Workflow Updates

Sharon Settnek reviewed the five workflow lines which together constitute the Prospective/Retrospective Biomarker Study Workflow (Attachment 2). No changes were made to this workflow. However, a discussion ensued regarding the governance of human tissues. Mark Rubin inquired whether any governance rules should be incorporated into the Biomarker Study workflow. He further asked whether user access to different types of data would have to be defined through the Administrative Core. Dr. Rubin concluded that any data that goes back to the Administrative Core would have to have a governance component. Paul Fearn suggested that the data elements submitted to the Administrative Core should be setup when the site registers the protocol. Bruce Trock agreed that this issue must be addressed, perhaps at the executive committee level (i.e., by the principal investigators [PI]) of the Prostate SPOREs.

Julie Schneider asked Sharon Settnek whether the issue of assigning specific levels of access to specific users would be addressed in the use case discussion. Sharon Settnek confirmed her statement and added that, once the use cases are completed, different rules (e.g., viewing, submission, querying, publishing) would be setup to determine what data will be available. She then asked Paul Fearn if there were any general governance rules that could be used in the use case sample assumptions. Paul Fearn replied that some general governance rules were detailed in

the request for proposals (RFP) for this project. Finally, Bruce Trock made the observation that the information for these use cases would have to be completed by the Prostate SPORE PIs when they meet to discuss how information will be shared.

Discussion of Next Steps

As the meeting came to a close, Sharon Settnek remarked that significant progress had been made on the workflow diagrams. She also informed participants that updated workflow diagrams would be provided to them in short order and suggested establishing a LISTSERV to facilitate the receipt of future workflow updates by the group. She then reviewed the next steps as follows.

- Before the next teleconference, all group members should:
    - Review the use case example completed by Dr. Peter Scardino in preparation for developing use cases.
    - Review the CDEs to be provided by Mariana González del Riego in preparation for the use cases.
- The next teleconference will be held on May 13, 2005 at 9:00 a.m. EDT. Mariana González del Riego will be providing additional details to participants via e-mail.
    - The focus of the next teleconference will be the completion of several high-level use cases and the derivation of associated CDEs.
- Items to be completed after the next teleconference call include the following:
    - Prioritization of use cases
    - Creation of a matrix of potential CDEs based on (1) associated use case subheadings, (2) CDEs needed, and (3) prostate cancer CDEs already available in the Cancer Data Standards Repository (caDSR), including Prostate SPORE CDEs.

The teleconference was officially adjourned at 9:30 a.m. EDT.

Attachment 1

caBIG/NBN Pilot Teleconference: Review of NBN Pilot Workflows and High-Level Use Cases
AGENDA

Friday • May 6, 2005
8:00 AM – 9:30 AM EDT

8:00 AM – 8:05 AM **Introductions**

- Participant Introductions
- Discuss Meeting Goals/Objectives
  - Review of First Iteration Workflows
  - Review and Further Development of First Iteration High-Level Use Cases
  - Understand Use Case Priorities to Assist in Defining Project Scope and Timelines
  - Discuss CDEs associated with Prioritized Use Cases

8:05 AM – 8:30 AM **Review First Iteration Workflows**

- Prospective Workflow
- Retrospective Workflow
- Biomarker Workflow

8:30 AM – 9:15 AM Discuss High-Level Use Cases

- Review Draft Use Cases
- Continue Drafting High-Level Use Cases

9:15 AM – 9:25 AM **Prioritize High-Level Use Cases**

9:25 AM – 9:30 AM **Discuss Next Steps**

- Scope Document Based on Prioritized Use Cases
- Review of Existing CDEs
- Mapping of Existing CDEs to Prioritized Use Cases

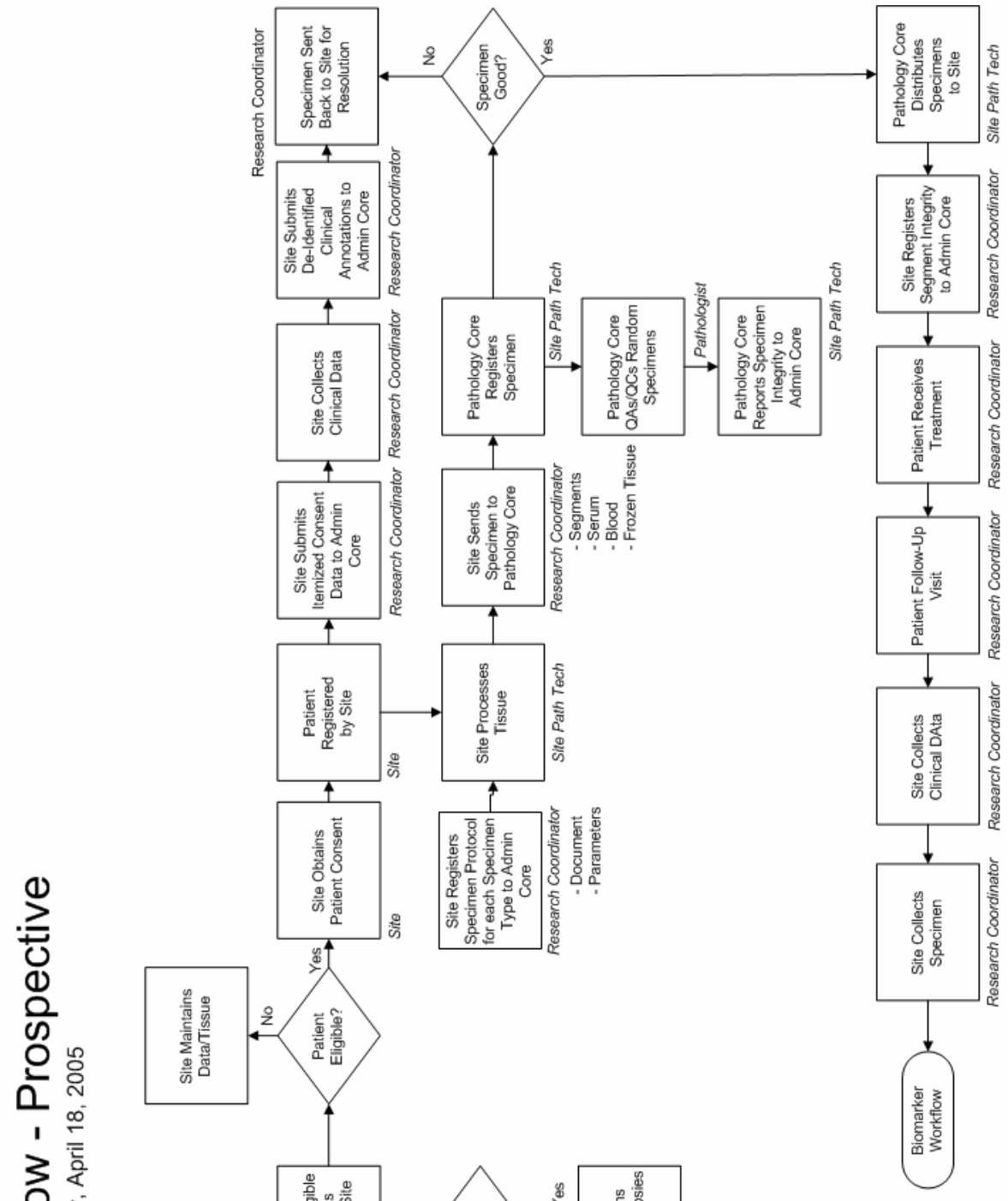
Additional Information

Draft Workflow Diagram and High-Level Use Case:

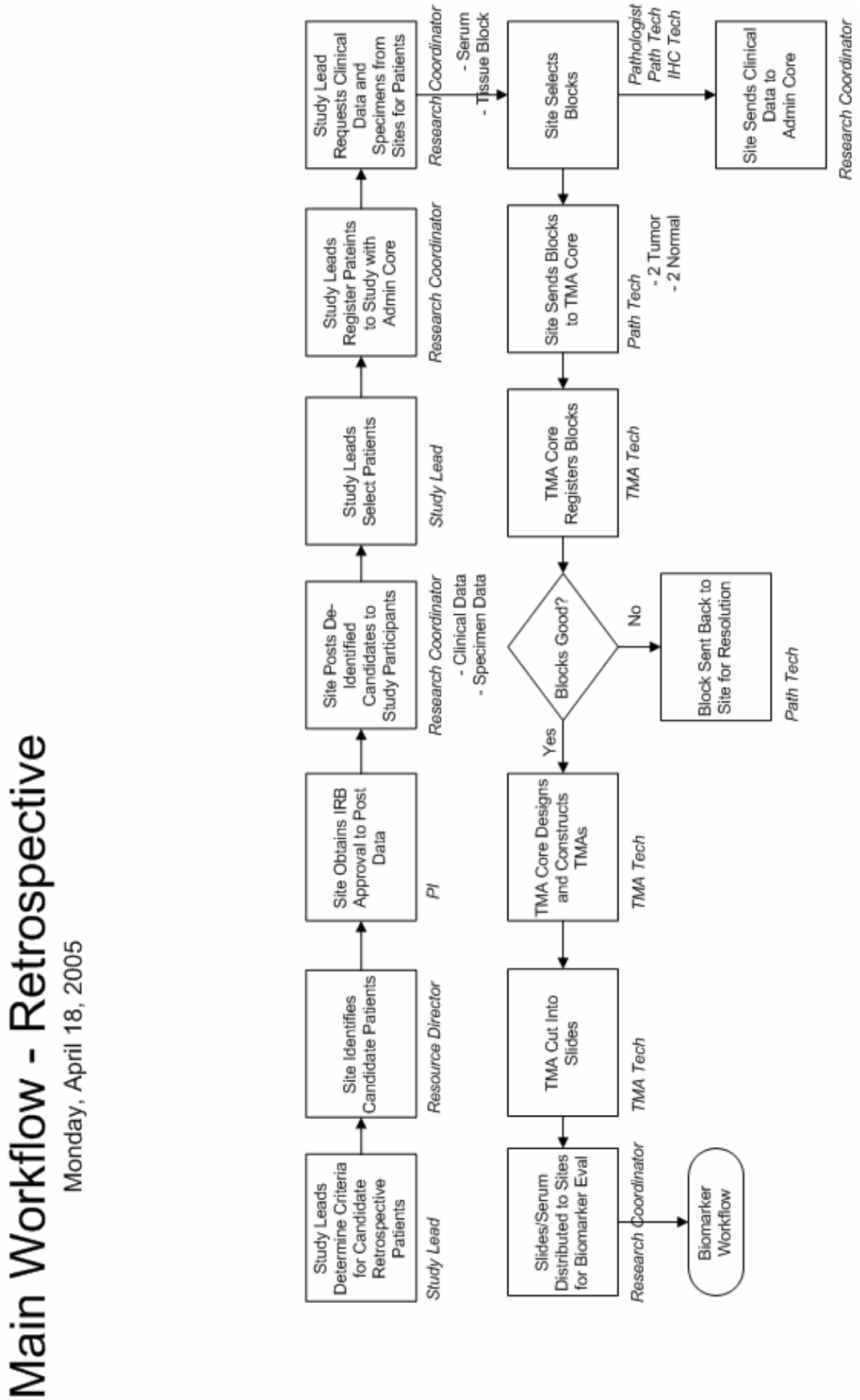
Prostate SPORE pathology CDEs: http://cdebrowser.nci.nih.gov/CDEBrowser/

(Note: On the left-hand side navigation bar, click on: CTEP → Protocol Form Template → Disease → Prostate. The pathology forms appear to be the most specific to the Prostate SPOREs.)

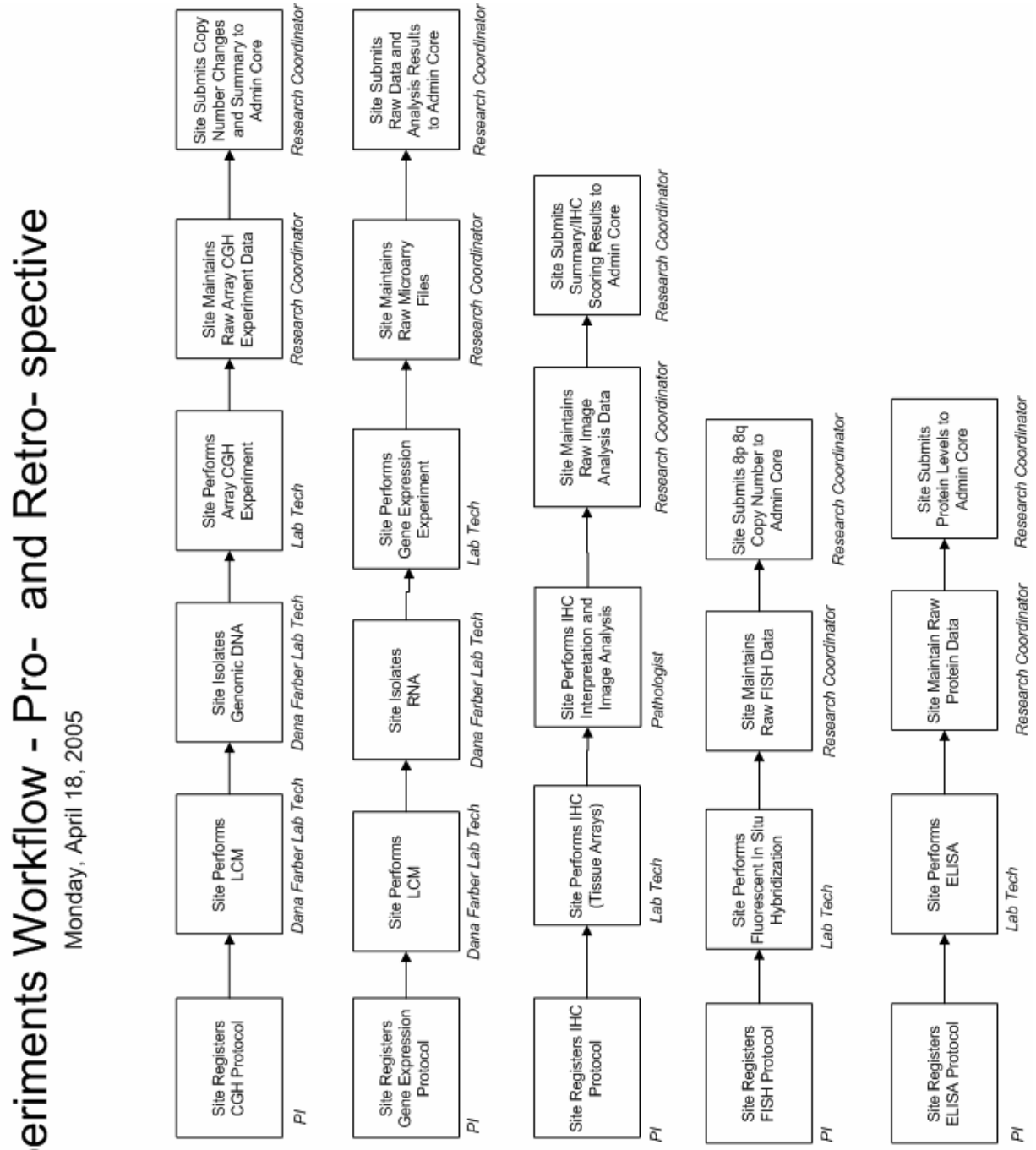
Attachment 2
Original Workflow Diagrams



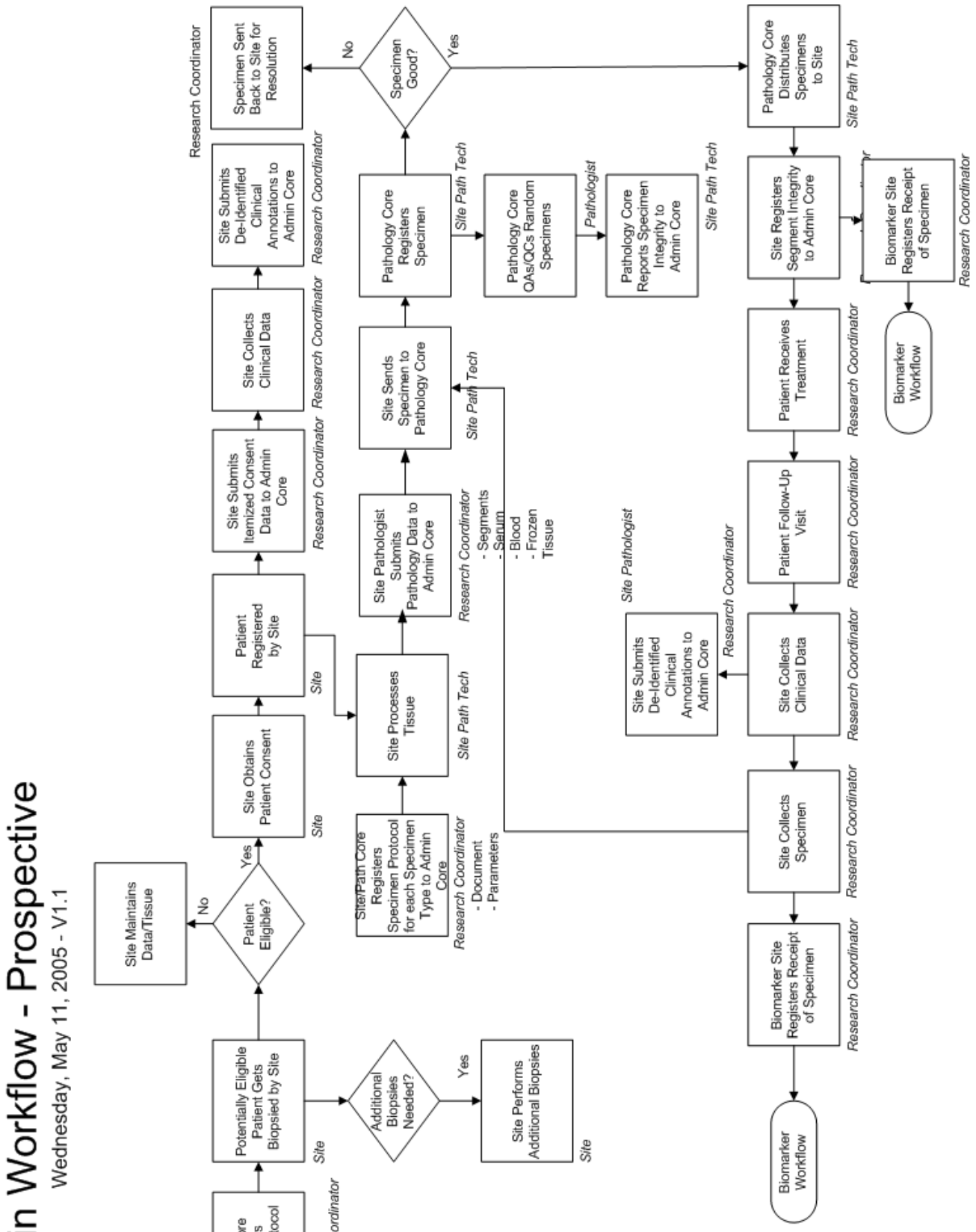
Attachment 2
Original Workflow Diagrams



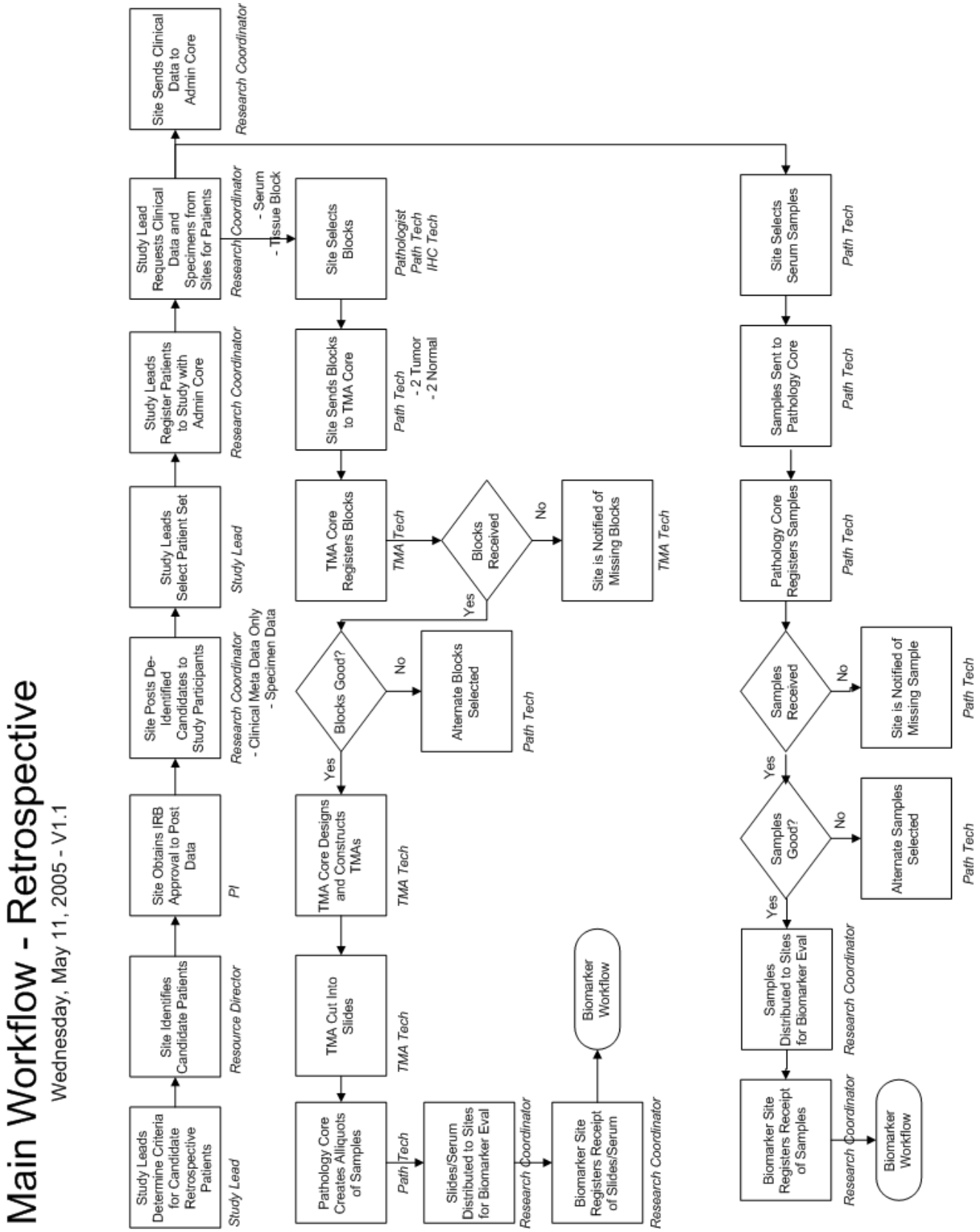
Attachment 2
Original Workflow Diagrams



Attachment 3
Updated Workflow Diagrams



Attachment 3
Updated Workflow Diagrams



Attachment 3
Updated Workflow Diagrams