Ancillary Study Management Systems: What Do We Need?

Elizabeth K. Nelson[1], Britt Piehler[1], Adam Rauch[1], Sarah Ramsay[2], Drienna Holman[2], Smita Asare[3,4], Adam Asare[3,4], Mark Igra[1]

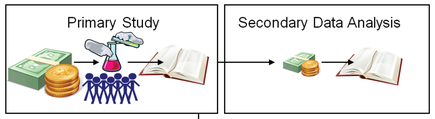
[1]*LabKey Software, Seattle, WA USA*
[2]*Statistical Center for HIV/AIDS Research & Prevention (SCHARP), Fred Hutchinson Cancer Research Center, Seattle, WA, USA*
[3]*University of California, San Francisco, CA, USA*
[4]*Immune Tolerance Network, Bethesda, MD, USA*

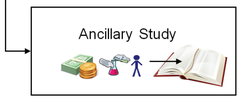
LabKey Software™
PARTNERS IN SCIENCE

Abstract

Ancillary studies allow researchers to leverage the high-value data collected as part of a primary trial or study, augment it with new measurements, and answer questions that were not part of the primary study design. Management of ancillary studies is complex, so tool support is desirable. We review the core scenarios that an ancillary study management system (ASMS) would need to support. We are currently developing an open-source, general-purpose ASMS based on LabKey Server [1] (http://www.labkey.org).

Introduction





The valuable clinical data, specimens and assay results collected during a primary clinical trial or observational study can enable researchers to to answer additional, pressing questions with relatively small investments in new measurements. However, management of such follow-on, "ancillary" studies is complex, typically requiring coordination across institutions, sites, repositories and approval boards, plus distributing, integrating and analyzing diverse data types. Surprisingly, general-purpose software systems that could simplify the management of ancillary studies have not yet been explored in the research literature [2].

Methods

We identified requirements for ancillary study management primarily as part of our ongoing work with a number of large research consortia. These organizations include the Center for HIV/AIDS Vaccine Immunology (CHAVI), the Immune Tolerance Network (ITN), the HIV Vaccine Trials Network (HVTN), and the U.S. Military HIV Research Program (MHRP). We also consulted with researchers at a range of disease research organizations regarding their workflows and data management strategies. Lastly, to enhance breadth, we reviewed process documents for ancillary study management from a variety of other organizations.

Results

Differentiating Characteristics of Ancillary Studies

(1) Collection of additional measurements
(2) Conduct of a study by external investigators
(3) Cross-protocol data pooling and analysis
(4) Pre-existing and new participant consent
(5) Pre-existing data context and provenance

Existing Alternatives for Managing Ancillary Studies

(1) Using a clinical trial management system (CTMS) for the aspects of ancillary studies that the tool can support
(2) Using an *ad hoc* combination of software for project management, data management and specimen requests and tracking
(3) Developing custom, organization-specific, end-to-end systems.

Key Scenarios for Ancillary Study Management

The steps listed in Table 1 provide a representative (but certainly not universal) workflow for ancillary study management. To focus the discussion, we consider a "freezer study" workflow, where existing specimens undergo additional analysis, but no new specimens or clinical data are collected. Our collaborators are most commonly concerned with this kind of ancillary study. This type of ancillary study is particularly illustrative of how the requirements for managing ancillary studies extend beyond those of either primary studies or secondary data analysis.

Any organization executing an ancillary study would have its own unique workflow, using these steps to varying degrees and in varying order, so an ASMS would need to be sufficiently flexible to accommodate this variability. Nevertheless, our review revealed surprising consistency in scenarios, likely due to the uniformity of regulatory funding requirements and published best practices.

Step	Key Differentiating Requirements
1. Hypothesis generation	a. Identifying interesting categories of participants from primary study b. Access existing data and specimen information for those participants c. Provide sufficient information to external investigators for them to propose ancillary studies
2. Proposal review	a. Review availability of existing specimens b. Review priorities for use of remaining specimens, possibly reserving specimens c. Evaluate overlap and comparability of existing data and proposed measurements/analyses
3. Creation of protocol or plan	a. Decide which existing participant data and specimens to use b. Plan expectations for collecting new ancillary data complementary to existing primary data
4. Verifying or updating consent	a. Determine whether consent exists and is sufficient for desired analyses b. Obtain additional consent if necessary
5. Retrieval of existing data	a. Compile relevant subset of primary study data required for the ancillary study b. If external investigators are leading the ancillary study, share this subset of data with them
6. Delivery and analysis of specimens	a. Request, locate, ship and track existing specimens b. Maintain identifiers relevant to primary study during further specimen analysis
7. Data integration	a. Retain context and provenance from both primary and ancillary studies, including processing and quality control information b. Retain origination information (primary vs. ancillary) c. Resolve differences (representation, quality control, etc) and join ancillary and primary data
8. Data/specimen repatriation	a. Contribute ancillary study data (raw and/or processed) back to the primary study b. Retain data context, provenance, processing and other metadata from ancillary study c. Return unused specimens
9. Publication	a. Coordinate preparation and review of publications across primary and ancillary investigators

Table 1: Key differentiating requirements for an ancillary study workflow. Note that this workflow addresses the common "freezer study" scenario, where the ancillary study's additional measurements come from analysis of existing specimens, not collection of new specimens or clinical data.

Software Development

In collaboration with CHAVI and the ITN, we are currently developing a general-purpose, open-source ASMS based on the LabKey Server system. LabKey Server provides support for managing, integrating, analyzing and sharing a wide range of experimental and clinical data types. The system already includes many features particularly helpful for ancillary study data management, including tools for:

(1) Sorting, filtering, aligning, and visualizing data
(2) Managing experimental data types
(3) Integrating varied types of data from different sources
(4) Marking columns with ontological concepts to indicate meanings and relationships
(5) Summarizing and reviewing available data (see Figure 1)

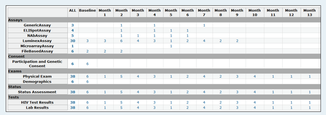


Figure 1: The study navigator lists all datasets the user has permission to view.

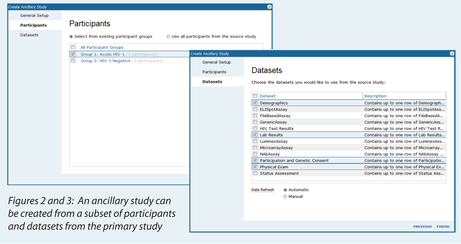
The system also already includes many features useful for operational management of ancillary studies, including:

(1) Role- and group-based security
(2) Specimen tracking and requests
(3) Issue tracker
(4) Wikis
(5) Message boards
(6) File management tools

Current and future development efforts address five key ancillary study scenarios:

(1) Sub-setting, tracking, and pooling participants within studies (see Figures 2 and 3)
(2) Reporting and analysis at the participant group level
(3) Scheduling and tracking expectations
(4) Providing data context
(5) Tracing data provenance

The current LabKey Server release already provides support for portions of scenarios (1)-(4).



Figures 2 and 3: An ancillary study can be created from a subset of participants and datasets from the primary study

Summary of Conclusions

The scenarios and requirements we describe can help guide the development of systems that can make conducting ancillary studies easier, less expensive and less error-prone. Given the relatively consistent characteristics and challenges of ancillary study management, general-purpose ASMSs are likely to be useful to a wide range of organizations.

Open-source, general-purpose ASMSs are particularly desirable due to their lack of licensing fees and openness to customization by individual organizations. Using the scenarios identified as part of this review, we are currently developing an open-source, general-purpose ASMS based on LabKey Server in collaboration with CHAVI and the ITN.

Get LabKey Server

Installers, source code, and documentation for LabKey Server are freely available: www.labkey.org

LabKey Server is maintained by a team of professional software engineers: www.labkey.com

References

1) Nelson E, Piehler B, Rauch A, *et al.* **Ancillary Study Management Systems: Review of Needs and Options.** *BMC Medical Informatics and Decision Making*. Submitted, 2012.

2) Nelson E, Piehler B, Eckels J, *et al.* **LabKey Server: An open source platform for scientific data integration, analysis and collaboration.** *BMC Bioinformatics*. 2011;12:71.

Acknowledgements

This work was supported by AI067854 Center for HIV/AIDS Vaccine Immunology (CHAVI) grant from NIAID/NIH, Division of AIDS; grant 38744 from The Bill and Melinda Gates Foundation to the Vaccine Immunology Statistical Center (VISC) and the CAVD; grant AI068635 from the NIAID to HVTN; contract #694251 from the Henry Jackson Foundation for MHRP to SCHARP; and contract #N01-AI15416 from NIAID/NIH to the ITN; and funds provided by the Juvenile Diabetes Research Foundation to ITN.