2019

U.S. FOOD & DRUG ADMINISTRATION

Source Data Capture from EHRs: Using Standardized Clinical Research Data

Mitra Rocca[1], Adam Asare[2,3], Laura Esserman[2], Sue Dubman[2], Gideon Gordon[1]

[1]Center for Drug Evaluation and Research, Food and Drug Administration
[2]University of California - San Francisco (UCSF)
[3] Quantum Leap Healthcare Collaborative

This project was funded by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (PCORTF) under the Intra-Departmental Delegation of Authority Request #075-X-0145-000 with the FDA.

This report reflects the views of the authors and should not be construed to represent FDA's views or policies.

October 30, 2019

TABLE OF CONTENTS

GLOSSARY AND ABBREVIATIONS

ACA	Affordable Care Act of 2010
API	Application Programming Interface
ASPE	Assistant Secretary for Planning and Evaluation
AWS	Amazon Web Services
BLA	Biologics License Application
BrCa	Breast cancer
caDSR	Cancer Data Standards Registry and Repository
CCD	Continuity of Care Document
CDA	Clinical Document Architecture
CDASH	Clinical Data Acquisition Standards Harmonization Project
CDC	Centers for Disease Control and Prevention
CDISC	Clinical Data Interchange Standards Consortium
CERSI	Centers of Excellence for Regulatory Science and Innovation
CFR	Code of Federal Regulations
CMS	Centers for Medicare and Medicaid Services (CMS)
CRF	Case Report Form
CSUCI	Computerized Systems Used in Clinical Investigations
CTCAE	Common Terminology Criteria for Adverse Events
CTMS	Clinical Trial Management System
DoD	Department of Defense (DoD)
eCOA	Electronic Clinical Outcomes Assessment (COA): a tool for capturing both patient and clinician reported outcomes
EDC	Electronic Data Capture
EHR	Electronic Health Record
EORTC	European Organisation for Research and Treatment of Cancer
ePRO	Electronic Patient Reported Outcomes
eSource	Electronic Source
FDA	Food and Drug Administration
FD&C Act	FDA Drug and Cosmetic Act
FHIR	Fast Healthcare Interoperability Resources
Health IT	Health Information Technology
HHS	Department of Health and Human Services
HL7	Health Level Seven
IHE	Integrating the Healthcare Enterprise
IND	Investigational New Drug Application

I-SPY 2 TRIAL	Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging And molecular 2
IVRS	Interactive Voice Response Systems
IWRS	Interactive Web Response Systems
LOINC	Logical Observation Identifiers Names and Codes
MedDRA	Medical Dictionary for Regulatory Activities. A global standard medical terminology designed to supersede other terminologies (such as COSTART and ICD9) used in the medical product development process.
MUGA	multigated acquisition
NCI	National Cancer Institute
NDA	New Drug Application
ODM	Operational Data Model
ONC	Office of the National Coordinator for Health Information Technology
PCOR	Patient-centered Outcomes Research
PCORTF	Patient-centered Outcomes Research Trust Fund
PROMIS	Patient Reported Outcomes Measurement Information System
PROPr	PROMIS-Preference
QoL	Quality of Life
RFD	Retrieve Form for Data capture
RIS	Radiology Information System
RxNorm	RxNorm is a normalized naming system for generic and branded drugs; and a tool for supporting semantic interoperation between drug terminologies and pharmacy knowledge base systems
RWD	Real World Data
RWE	Real World Evidence
SDV	Source Data Verification
SOAP	Simple Object Access Protocol
SCDM	Society for Clinical Data Management
SDTM	Study Data Tabulation Model
SNOMED CT	Systemized Nomenclature in Medicine – Clinical Terminology
TAUG	Therapeutic Area User Guide
UAMS	University of Arkansas for Medical Sciences
UCSF	University of California in San Francisco
USCDI	US Core Data for Interoperability
VA	U.S. Department of Veterans Affairs

EXECUTIVE SUMMARY

Accessing research data directly from Electronic Health Records (EHRs), known as electronic source data capture (eSource), can create efficiencies in the clinical research process while improving data quality, reducing cost, maintaining integrity and preserving audit trails. A significant portion of the growing costs of clinical trials, and hence drug development, relates to source data verification (SDV), a process by which data from clinical trial collection systems are compared to the source information. The use of Electronic Health Records (EHRs) in clinical research has the potential to eliminate the need for this comparison, and for this reason electronic source data capture (eSource) from EHRs has been a priority for the US Food and Drug Administration (FDA) and the subject of a guidance published in 2018.[1]

The benefits of Electronic Source data capture (eSource) include: 1) decreasing the burden on healthcare providers and research staff in conducting research; 2) improving the quality of data submitted for regulatory decision-making; and 3) allowing for more efficient use and re-use of healthcare data to support high quality clinical care delivery and clinical research participation. Our research makes clear that accomplishing these goals will require solutions that go beyond electronic data transfer and address certain fundamental issues, namely 1) data representation in EHRs, 2) heterogeneity between healthcare and clinical research data requirements, and 3) a need to streamline clinical practice processes to support clinical research. Until these issues are addressed, eSource will be an effective and important, but only partial, solution to the problems it seeks to resolve.

The OneSource Project, a collaboration between investigators at the University of California San Francisco (UCSF) and FDA, was established with the goal of developing methods and tools to automate the flow of structured EHR data into external systems and thereby reduce operating costs, save time, and improve data quality for clinical trials. We have demonstrated an approach to transmit structured data from the UCSF EHR system to a clinical trial electronic data capture (EDC) system. In this approach, we populated Electronic Case Report Forms (eCRFs) for a phase II clinical trial, (**I**nvestigation of **S**erial Studies to **P**redict **Y**our **T**herapeutic **R**esponse with **I**maging **A**nd mo**L**ecular Analysis **2** (I-SPY 2 TRIAL)). OneSource leveraged standards from Health Level 7 (HL7), Clinical Data Interchange Standards Consortium (CDISC), and Integrating the Healthcare Enterprise (IHE) for the capture and transmission of clinical research data. The goal of the work described here was to harmonize the data elements and use better tools for data capture that could make these key elements available for healthcare providers. Electronic Patient Reported Outcomes (ePRO) was also implemented as part of OneSource for source data capture from patients in the I-SPY 2 TRIAL.

This report describes the development process and approach to electronic source implementation at UCSF. It includes publicly available resources that can be leveraged by the research community. We suggest that the approach taken in this project could become a model for effective, efficient use of clinical data for clinical research and decision-making, leading to cost reduction, time savings, and improved data quality for future clinical trials.

BACKGROUND

The current process by which data is collected in clinical trial systems generally diverges from health care systems, which translates into serious inefficiencies in the conduct of clinical trials.

The parties that consume clinical data all need the same information - a consistent record where the diagnosis is described accurately and completely; the events of clinical care are captured with fidelity; quality-of-life impact is recorded, assessed and visible; and the follow-up (both adverse events and recurrence information) is complete and up-to-date. If we converged on a system where the key elements were captured accurately at the point of care, and in a manner that facilitated secondary use, not only would healthcare providers have them at their fingertips and be better equipped to deliver appropriate care and interventions, but clinical trials would become efficient.

Clinical trialists often say that the data required for trial participation is better and cleaner than for clinical care. However, before a healthcare provider has a cancer procedure or chemotherapy there is a need for high-fidelity data and a checklist of mission-critical data to ensure that good clinical decisions are made. Healthcare providers and patients need better, more efficient systems that are consistent and standardized and allow healthcare providers to focus on the capture of high-quality data that are distilled to the mission-critical elements that will guide care. Such systems would also make high-quality data available for clinical research, registries, and quality improvement, and avoid the redundancies and occasional inconsistencies that reside between EDC and EHR systems.[2]

Data standards are needed for the collection of patient-reported data such as quality-of-life assessments and adverse event reporting. It is also important that the data flow between healthcare and clinical research systems is bidirectional. When a patient participates in a clinical trial, the data generated by the trial, or at least the summary of the trial data, must be reported back to the healthcare provider and patient for consideration in their on-going care. A clinical trial summary for each patient would add value to the health care providers and vice versa.

The project described here (The Source Data Capture from Electronic Health Records (EHRs): Using Standardized Clinical Research Data Project (OneSource)) has created a framework for using EHR data as Electronic Source (eSource) in clinical trials that support medical product applications[i] with the goal of increasing trial efficiency and reducing costs. This effort is a collaboration between FDA and the University of California - San Francisco (UCSF) which, along with Stanford University, serves as one of the FDA Centers of Excellence for Regulatory Science and Innovation (CERSI).[3] The Open-source forms, source code, and standards enhancement recommendations developed in the course of this project are being released to the public and organizations interested in using EHRs for conducting clinical research.

The clinical data management environment in OneSource is secure and conforms to 1) federal regulations:

[i] Specifically, New Drug Applications (INDs), New Drug Application (NDAs) or Biologics License Applications (BLAs).

Part 11 of Title 21 of the Code of Federal Regulations (21 CFR Part 11), defining criteria under which electronic records and electronic signatures are considered trustworthy, reliable, and equivalent to paper records[4] and 2) the FDA guidance for industry *Computerized Systems Used in Clinical Investigations (CSUSI)*, which is intended to enhance the reliability, quality, and integrity of electronic source data and source documentation.[5] Furthermore, OneSource leverages open, consensus-based standards (e.g., CDISC, HL7, IHE).

OneSource focuses on collecting data elements required for I-SPY 2 TRIAL, which consists of large adaptive clinical trials that simultaneously tests drug treatments for breast cancer using biomarkers and collects Electronic Patient Reported Outcomes (ePROs). A series of Case Report Forms (CRFs) and electronic patient reported (ePRO) forms were designed and implemented within the UCSF EHR system test environment to collect data for the I-SPY 2 clinical trials in a structured and standardized fashion and to populate the healthcare provider dashboard CRFs from the UCSF EHR system.

The electronic capture of data from EHRs and healthcare devices such as Electronic Patient-Reported Outcomes (ePRO) devices, digital imaging, and mobile health devices could improve the reliability, quality, traceability (provenance), and integrity of data from electronic source to regulatory submission. In pursuing this goal, OneSource follows recommendations in two FDA guidances: 1) "Electronic Source Data in Clinical Investigations" (eSource), which encourages use of electronic source data in the conduct of regulated clinical trials and 2) "Use of Electronic Health Record Data in Clinical investigations," which encourages use of EHRs in FDA-regulated clinical investigations and promotes the interoperability of EHR and Electronic Data Capture (EDC) systems.[6] OneSource provides a working example of the eSource approach[1] and produces guidelines that could be used by other researchers to facilitate implementation.[ii]

The funding for this project was provided by the Patient-Centered Outcomes Research Trust Fund through a competitive application process administered by the Associate Secretary for Planning and Evaluation (ASPE) of the Department of Health and Human Services.[7]

[ii] In addition to these two guidance documents, the passage of the 21st Century Cures Act has been critical to accelerate medical product development and bring new innovations and advances faster and more efficiently to the patients who need them. Under the Cures Act, FDA created a framework for evaluating the potential use of Real World Data (RWD) to generate Real World Evidence (RWE) of product effectiveness to help support approval of new indications for drugs approved under FDA Drug and Cosmetic (FD&C) Act Section 505(c) or to help to support or satisfy post approval study requirements.

INTRODUCTION

Goals and Objectives

The goal of this project is to demonstrate a method of transferring specific data, or health information, from a patient's electronic health record (EHR) to an electronic data capture (EDC) system for collecting clinical trials data. If this project is to be useful to all researchers, the proposed solution must use open, consensus-based standards that dictate the structure and format of the data.

The project's objectives are to:

1. Demonstrate an end-to-end (EHR to EDC) standards-based technology solution for the capture and transmission of regulated clinical research data by leveraging the following resources:
    - Health Level Seven (HL7) Continuity of Care Document (CCD)
    - Integrating the Healthcare Enterprise (IHE) Retrieve Form for Data Capture (RFD)
    - HL7 Fast Healthcare Interoperability Resources (FHIR), and
    - Clinical Data Interchange Standards Consortium (CDISC) standards

2. Assess the utility of the standards-based technology solution processes for FDA inspection and reconstruction of clinical investigations
3. Develop guidelines for future implementations in both healthcare and clinical research
4. Provide recommendations for the improvement of existing standards and implementations
5. Develop a general framework (technologies, processes, policies, governance and standards) for the electronic source data capture systems in regulated clinical trials and electronic patient reported outcomes (ePRO)

Problem Statement

The information systems and the underlying data models and standards that define clinical care and regulated clinical research are highly variable. This lack of uniformity was not an issue for the conduct of regulated clinical research prior to use of EHRs or EDCs, because data were captured on paper case report forms. However, much has changed in the past decade for regulated clinical research where EDC systems are now ubiquitous for the capture of clinical trials data. Similarly, EHRs and other Health Information Technology (Health IT) systems have been widely adopted and are rapidly becoming a standard part of clinical care. Today, most hospitals and health care providers in the US have a digital footprint. As of 2015, 96 percent of nonfederal acute care hospitals and 78 percent of office-based physicians adopted certified health IT.[8] The increase in adoption of health IT means that most Americans receiving health care services now have their health data recorded electronically. However, this information is not easily accessible to clinical researchers.

Structured electronic data capture (EDC) used in clinical trials enables the collection of high-fidelity, usable information, but the effort is almost completely manual. The principal clinical source data is most often the EHR, which introduces significant costs in terms of data processing and cleaning. Over 70% of data are duplicated between an institution's EHR and clinical trial systems.[9,10]

The data that is organized and collected for the clinical trial is the very data that would streamline and improve clinical care. This highlights the need for tools that allow collection of data that is interoperable and prevents patients from having to report the same data multiple times.

Solution and Implementation

Phase One of the OneSource Project demonstrated an approach and developed a framework for collecting data for clinical trials that populates an Electronic Data Capture (EDC) system directly from an EHR system. The OneSource project also supports electronic source data capture from patients using an electronic Patient Reported Outcomes (ePRO) platform that is integrated with source data capture from EHRs for clinical investigations (see Figure 1). This approach is designed for FDA-regulated clinical investigations using open, consensus-based data standards used in health care and clinical research (e.g., Health Level Seven (HL7), Clinical Data Interchange Standards Consortium (CDISC)) as well as integration profiles (e.g., Integrating the Healthcare Enterprise (IHE) Retrieve Form Data capture (RFD) profile). The IHE RFD integration profile enables a clinical investigator to display a partially completed case report form within the EHR system. RFD provides a method to collect data from within the EHR application in a way that will meet the requirements of an external system (in this case the EDC system).[11]

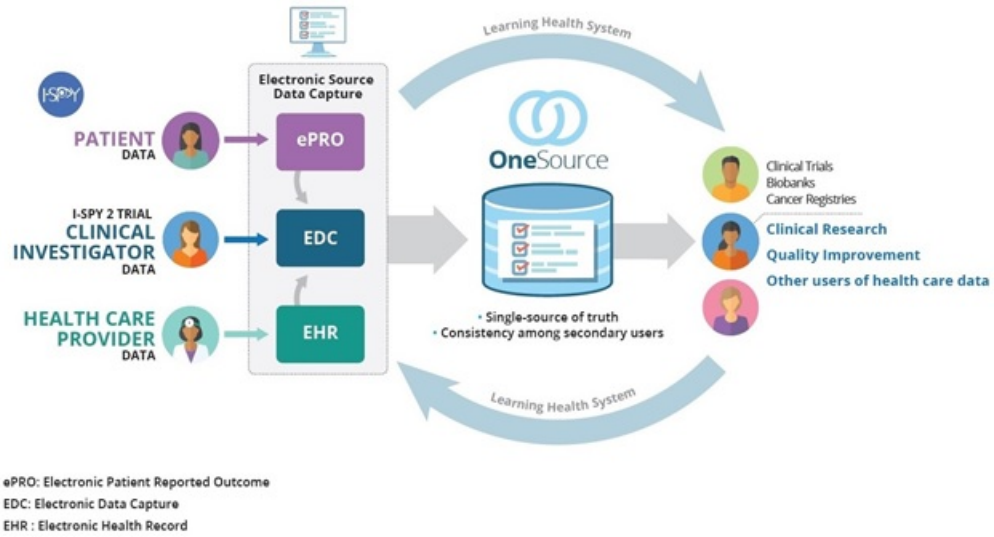


Figure 1: Overview of OneSource electronic source data capture system supporting clinical care and clinical investigations.

The key deliverables of this project are the following:

- I-SPY 2 TRIAL Case Report Forms (CRFs) specific to eSource data capture: The CRFs include CDASH mappings that can be leveraged by researchers implementing clinical studies that have standardized terms for data interoperability.
- Source code for EHR integration: The source code can be used by researchers and EHR implementation specialists that wish to automate EHR data capture for insertion in study CRFs.
- Gap analysis between the EHR and the I-SPY 2 case report forms: The gap analysis demonstrated the percentage and feasibility of capturing discrete data elements from an institution's EHR, and the amount of manual abstraction still required for manual entry.
- Electronic Patient Reported Outcomes (ePRO) forms for patient reported outcomes: The ePRO forms can be leveraged by clinical study researchers with standardized survey questions to support data interoperability and data sharing.

Deliverables have been placed in the public domain for interested organizations and PCOR researchers who may want to apply this approach to collect data from EHRs in their clinical research. (See appendices A - E).

METHODS

OneSource collects structured data for both clinical care and regulated clinical research. The UCSF EHR system is used by healthcare providers to collect and store the healthcare information of patients at the UCSF hospitals and clinics. Several standards are used to integrate the UCSF EHR system with the I-SPY 2 EDC system. Figure 2 is a high-level illustration of the EHR-EDC workflow. Currently, the clinical investigators need to enter data in the UCSF EHR system as well as the I-SPY 2 EDC system. In the future, data from the UCSF EHR system (e.g., healthcare provider notes) as well as data from other clinical information systems (e.g., Radiology Information Systems (RIS), Pathology Information System) will be transferred to the EDC system for the I-SPY 2 TRIAL clinical trial. In addition, the data elements from the UCSF EHR system will be mapped to the CDISC Therapeutic Area User Guide Data Standards for breast cancer (TAUG-BrCA).

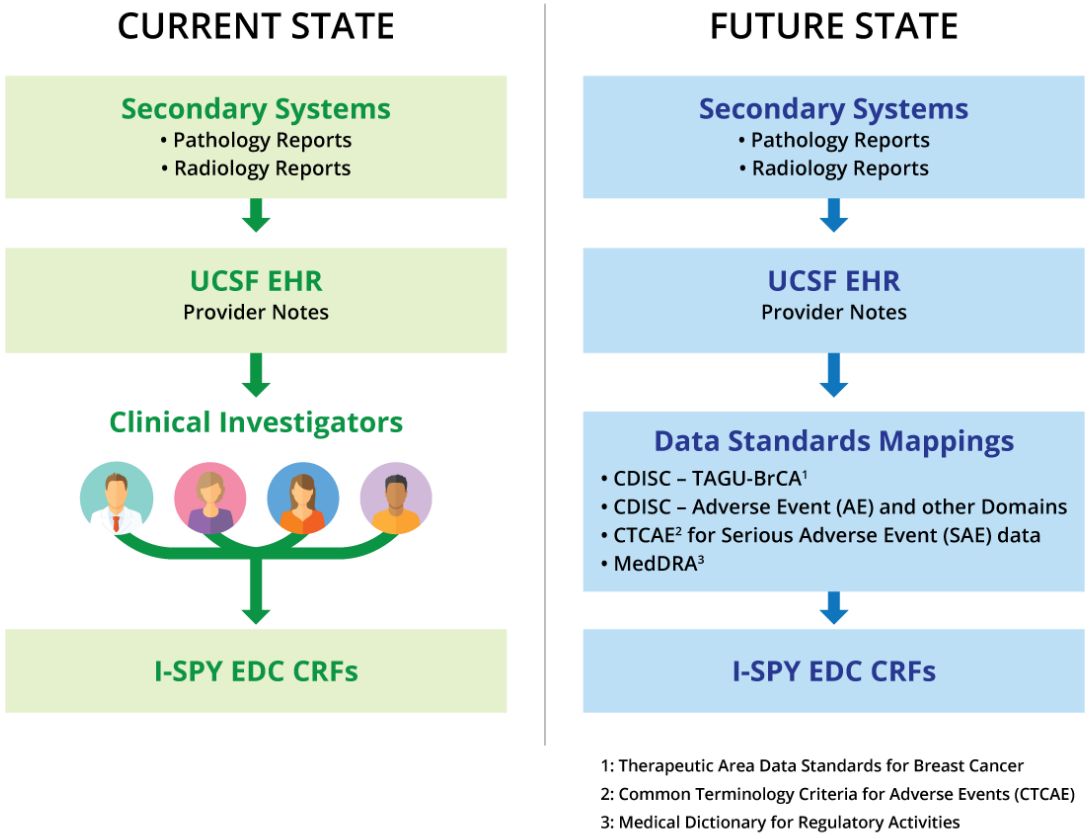


Figure 2: I-SPY 2 TRIAL eSource Electronic Data Capture Workflow: Current vs Future State

Technical Architectural Framework

In the OneSource Demonstration Project, data elements were captured using HL7 standards (e.g., HL7 CCD) from the UCSF EHR system as part of the clinical care process based on workflows for the clinical investigator. The EHR data elements pass through the OneSource data mapping framework and produce the CDISC clinical trial representation standard (e.g., CDISC ODM) used to populate the I-SPY 2 EDC system. Appendix A provides an overview of the OneSource components used for data exchange and the UCSF EHR system.

Project Constraints/Assumptions

This project had the following constraints and assumptions:

- Implementation was at a single site, UCSF.
- A prototype/demonstration project was tested as opposed to a system in a production environment.
- Implementation was with a single, proprietary EDC system.
- Implementation was done with a single, proprietary EHR system (based on UCSF's implementation of Epic).
- Content and work product developed is specific for a single therapeutic area, breast cancer.

Vendor Selection

A landscape assessment was performed to identify and evaluate clinical data management solutions, initially focusing on identifying electronic data capture (EDC) platforms suitable for electronic Patient Reported Outcomes (ePRO) and appropriate for integration into a comprehensive clinical research data management system. A list of vendors with established EDC platforms was developed based upon various in-depth research criteria. A series of questions was developed in collaboration with I-SPY 2 TRIAL staff to effectively and fairly evaluate each vendor (Appendix B).

After initial contact, a high-level version of the questions (main topics without the sub topics) was provided to vendors in advance; more in-depth issues were probed during the follow-up. Several of the questions were not interpreted uniformly by vendors. Additional research was conducted to confirm the architecture, customer base, and business strength of each EDC vendor.

Twenty-one potential vendors were identified through the research conducted by the UCSF team. Complete interviews and information were obtained for 11 EDC and ePRO vendors.

The project team at UCSF used the following key assessment criteria and business requirements to evaluate and select the EDC/ePRO vendor for this project:

1. Track record of success and clinical trials submission to the FDA for more than 10 years
2. Customer base
3. Ease in administration of the platform as a cloud-based solution
4. Capacity to support collection of electronic Patient Reported Outcomes (ePRO) data
5. Reporting capability and data integration of clinical care, clinical research and ePROs
6. Usability for providers, patients, clinical researchers, and others
7. Integration with the UCSF EHR and potentially EHR systems at other I-SPY TRIAL sites

Based on this assessment, the UCSF team selected OpenClinica as the ePRO vendor for OneSource and will be transitioning the EDC to OpenClinica for the next phase of the project.

Data Standards

The following data standards were used in the OneSource project:

- **HL7 Continuity of Care Document (CCD)**, a standard specifying the encoding, structure, and semantics of a patient summary clinical document for exchange.
- **Integrating the Healthcare Enterprise (IHE) Retrieve Form for Data Capture (RFD),** an integration profile that enables a clinical investigator to display a partially completed case report form within the EHR system. It is a method for retrieval of forms data from a forms source (in this project the UCSF EHR system) to meet the requirements of an external system (in this project the I-SPY 2 EDC system). (For the RFD technical implementation, see Appendix C.)
- **CDISC Operational Data Model (ODM)**, a vendor-neutral format for exchanging and archiving clinical research data, along with associated metadata, administrative data, reference data, and audit information. ODM has become the language of choice for representing, importing, and exporting case report forms.
- **Therapeutic Area Data Standards User Guide for Breast Cancer (TAUG-BrCa)**, which describes how to use CDISC standards to represent data pertaining to breast cancer studies. The focus of the TAUG-BrCa is on clinical trials of drugs to treat invasive breast cancer in neoadjuvant, adjuvant, and metastatic settings.

Implementation

Implementation of the system addressed two areas: the EHR/EDC system integration using IHE RFD and the ePRO questionnaire development.

EHR/EDC Systems Integration using IHE RFD

After the EDC system was selected, the UCSF team focused on reviewing the I-SPY 2 breast cancer case report forms, identifying and extracting data elements out of the UCSF EHR system. In addition, the team identified missing data elements in the UCSF EHR system (Appendix E), needed for I-SPY 2 breast cancer trials. The IHE RFD profile was leveraged to display the electronic Case Report Forms (eCRFs) within the EPIC EHR system and pre-populate forms.

ePRO Questionnaire Development

The I-SPY 2 clinical trial has approximately 1537 patients enrolled and randomly assigned across 18 clinical sites in the United States. Paper-based European Organization for Research and Treatment of Cancer (EORTC) Quality of Life (QoL) and Patient Reported Outcomes Measurement Information System (PROMIS®) Health Measures questionnaires have been distributed by paper for the first five years of this clinical trial study. The UCSF team deployed the OpenClinica Participate™ platform for ePRO surveys, which was provided to patients and accessible using mobile technology platforms.

Both PROMIS and Patient Reported Outcomes – Common Terminology Criteria for Adverse Events (PRO-CTCAE™) survey questions were administered to address the following Adverse Event (AE) areas: Anxiety, Depression, Sexual Function, and Sleep.

Please see Appendix F for ePRO surveys that were used in this project.

RESULTS

Accomplishments and Deliverables

Objective 1: Demonstrate an end-to-end (EHR to EDC) standards-based technology solution leveraging HL7 CCD, IHE RFD, HL7 FHIR and CDISC standards for the capture and transmission of regulated clinical research data

- **Deliverable:** Publicly available metrics on concordance of EDC CRFs with the EHR system shown in this report
- **Target Audience:** Biomedical informaticians, Health IT and EDC vendors, data standards developers, clinical researchers, and the biopharmaceutical industry
- **How it can be used:** Provides understanding to those implementing EDC systems for clinical trials of the level of concordance and 1:1 mapping of specific data elements in a clinical trial EDC system that matches to EHR data elements.
- **Access to Resource:** Source code for IHE RFD standard and EHR (EPIC Integration code base), available in Appendix D.

Objective 2: Assess the utility of the standards-based technology solution processes for FDA inspection and reconstruction of clinical investigations

- **Deliverable:** Gap analysis between clinical data elements collected in a healthcare setting by EHRs vs. clinical data elements required for regulated clinical research, focusing on three key I-SPY 2 TRIAL Case Report Forms as examples
- **Target Audience:** Biomedical Informaticians, EHR Implementation specialists, and CDISC, HL7, and IHE data standards experts
- **How it can be used:** Provides understanding to those implementing EDC systems for clinical trials of the level of concordance and 1:1 mapping of specific data elements in a clinical trial EDC system that matches to EHR data elements and the level of quality improvement that could be achieved through direct source capture of these elements

- **Access to Resource:** Detailed Gap analysis between CRFs and UCSF EHR system is in Appendix E.

Objective 3: Develop guidelines for future source data capture implementations in supporting both healthcare and clinical research

- **Deliverable:** Guidelines and recommendations in this report for use by PCOR researchers
- **Target Audience:** Biomedical Informaticians, EHR Implementation specialists, and CDISC, HL7, and IHE data standards experts
- **How it can be used:** Recommendations provided in this report can be leveraged for researchers interested in electronic source solutions. It will provide the researchers with a better and understanding of the technical challenges and the level of expertise needed.
- **Access to Resource:** See Report Discussion and Future Plans and Recommendations.

Objective 4: Provide recommendations for the improvement of existing standards and implementations

- **Deliverable:** Detailed recommendations in this report to Standards Development Organizations (e.g. CDISC, HL7 and IHE) to enhance their standards and implementations
- **Target Audience:** Biomedical Informaticians, EHR Implementation specialists, and CDISC, HL7 and IHE data standards experts
- **How it can be used:** Recommendations provided in this report can be leveraged for researchers interested in electronic source solutions. It will provide the researchers with a better and understanding of the technical challenges and the level of expertise needed.
- **Access to Resource:** See Report Discussion and Future Plans and Recommendations.

Objective 5: Develop a general framework (technologies, processes, policies, governance and standards) for the electronic source data capture systems in regulated clinical trials and electronic patient-reported outcomes.

- **Deliverables:** A general framework described in this report for electronic source data capture systems used in regulated clinical investigations and EDC/ePRO platforms assessment criteria to support this framework
- **Target Audience:** Biomedical Informaticians, EHR Implementation specialists, CDISC, HL7 and those involved in IHE data standards implementation
- **How it can be used:** Framework, code base, and forms can be leveraged by health IT systems integration specialists; data standards can be implemented by CDISC, HL7 and IHE data standard experts. **Access to Resource:** See current report and technical framework (Appendix A), IHE RFD SOAP codebase explanation (Appendix C), downloadable CRF/IHE RFD Codebase Resources, annotated CRFs with CDISC standards (Appendix D), and gap analysis between CRFs and UCSF EHR system (Appendix E).

Gap Analysis between EHR and EDC CRFs

A gap analysis was conducted between the Data Elements in UCSF's EHR system and I-SPY 2 TRIAL EDC Case Report Forms (CRFs) that had the highest number of discrete data elements mapping to the UCSF EHR system.

Based on the gap analysis and current CRF quality metrics, it was determined that having an eSource solution would dramatically improve data quality within the I-SPY 2 clinical trial program (see table 1). We evaluated "data rejection" reasons for the I-SPY 2 TRIAL "Baseline symptoms" CRFs over a 6-month period. The following criteria were used to evaluate quality metrics in conducting the gap analysis:

- Source Mismatch: I-SPY 2 TRIAL CRF data elements not matching to its corresponding field in the EHR
- Missing Source: hard copy source data missing on EHR printout or hand-written forms
- Redaction Issue: source data may have potential Protected Health Information (PHI) data that needs to be redacted and reloaded
- Incomplete Form: CRFs not filled out (In most instances, a direct eSource solution that mapped directly to the EHR would avoid the majority of these errors or missing information.)
- Upon Site Request: The site is aware of an issue in entry on their end, so sends an email request to monitors to record as rejection to allow revision by site personnel to update or add additional information that had been missed or incorrectly inserted
- Other: a catch all for all other issues

Table 1: I-SPY 2 TRIAL example CRF quality issue monitoring using standard manual entry approach for discrete field capture

Discordance issue/Flag	Proportion of data elements with issue
Source Mismatch	53.67%
Redaction Issue	12.21%
Missing Source	9.74%
Incomplete Form	4.87%
Upon Site Request	4.87%
Other	14.64%

Table 1 summarizes the reasons for rejection for these baseline symptoms prior to implementation of EHR/EDC integration.

We assessed concordance of I-SPY 2 TRIAL CRF data elements directly to EHR discrete data fields to determine the level of improvement in data quality and efficiencies that would occur with implementation of a source data capture solution of data directly from EHRs. The three I-SPY 2 TRIAL EDC CRFs with the highest number of CRF data variables that directly mapped to the EHR were "Laboratory and Test Results," "Menopausal Status," and "Baseline Data Elements."

Lab and Test Results

"Lab and Test Results" fields are typically well structured compared to most I-SPY 2 TRIAL CRFs. Table D1 in the Appendix E shows the mapping between the "Lab and Test Results" CRF data elements and the CDISC Clinical Data Acquisition Standards Harmonization (CDASH) standard. It also shows whether these data elements are currently collected in the UCSF EHR system.

Of the fifteen data elements in the "Lab and Test Results" CRF (Appendix E, Tables E1, E2 and E3), the following ten data elements (66%) are structured in Epic and map 1:1 with I-SPY 2 TRIAL data elements:

- Leukocytes
- Absolute Neutrophils
- Platelets
- Total Bilirubin
- Aspartate Aminotransferase (AST)
- Alanine Aminotransferase (ALT)
- Creatinine
- Cardiac Echo (Echocardiogram)
- Cardiac Echo Date of Procedure

- Cardiac Echo Left Ventricular Ejection Fraction (LVEF) %

Five of the fifteen data elements (33%) are inconsistently entered into the UCSF Epic EHR system. For example, either multigated acquisition (MUGA) Scan or Cardiac Echo may be present. Users may fill out one or the other using the "Transthoracic Echo" result. Please see detailed summary in Appendix E1.

In summary, with respect to the “Lab and Test” data elements: Ten (67%) are re-usable in the I-SPY 2 trial, and potentially all 15 would be if processes for entering this data were more clearly defined.

Menopausal Status

Tables E4, E5, and E6 in Appendix E illustrates the mapping between the “Menopausal Status” data elements in the I-SPY 2 Menopausal Status CFR data elements and the CDISC Clinical Data Acquisition Standards Harmonization (CDASH) standard, and whether these data elements are currently collected in the UCSF EHR system.

Of the twelve “Menopausal Status” fields in I-SPY 2 TRIAL, two of them, “Hysterectomy Date” and “Hysterectomy” are structured in the UCSF Epic EHR and map 1:1 with I-SPY 2 TRIAL data elements.

Baseline Data Elements

The Baseline CRF data elements are divided into 3 groups: Allergies, Baseline Condition, and Baseline Symptoms. Tables E7, E8, E9 and E10 in Appendix E illustrate the mappings between data elements in the I-SPY 2 TRIAL “Baseline” CRF data elements and the CDISC CDASH standard, and if these data elements are currently collected in the UCSF EHR system.

Seven (22%) of the 30 data elements needed for the Baseline CRF map directly to the UCSF EHR system data elements.

Use of standard based technology for Electronic Patient Reported Outcomes (ePRO)

Another component of the OneSource project and framework is implementation of source data capture from patients participating in the I-SPY 2 TRIAL. The I-SPY 2 TRIAL currently has patients enrolled and randomly assigned to 16 clinical sites in the United States. Survey instruments were distributed by paper to patients starting in January 2012 and included the following:

- The European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30, incorporating nine multi-item scales: five functional scales (physical, role, cognitive, emotional, and social); three symptom scales (fatigue, pain, and nausea and vomiting); and a global health and quality-of-life scale. as well as several single-item symptom measures.
- The EORTC QLQ-BR23, a breast-specific module consisting of 23 questions to assess body image, sexual functioning, sexual enjoyment, future perspective, systemic therapy side effects, breast symptoms, arm symptoms, and upset by hair loss
- Patient-Reported Outcomes Measurement Information System® (PROMIS), a set of person-centered measures that evaluates and monitors physical, mental, and social health in adults and children and can be used with the general population and with individuals living with chronic conditions

A decision was made to move from paper to ePRO in 2018. In the ePRO release, PROMIS measures selected addressed the following areas: Anxiety, Depression, Fatigue, Physical function, Sexual Function, and Sleep. For adverse events, the National Cancer Institute had developed a patient-reported outcomes version of the Common Terminology Criteria for Adverse Events **(**CTCAE™) termed PRO-CTCAE™, which consists of a series of questions that characterize the frequency, severity, and/or interference of 78 symptomatic treatment toxicities across 14 domains. The PRO-CTCAE are focused on physical functions, including symptomatic toxicities such as pain, fatigue, nausea, and cutaneous side effects such as rash and hand-foot syndrome. To reduce the survey burden on patients, I-SPY investigators, leadership, and patient advocates identified a set of PRO-CTCAE domains and items relevant to the population to include in the ePRO survey.

Beginning in July 2019, we deployed the revised ePRO surveys to the OpenClinica Participate™ test environment and leveraged OpenClinica Insight™ for alerts and periodic reports of ePRO results. The reports will be sent to healthcare providers and site coordinators.

Summary ePRO reports and longitudinal plot summaries (Figure 3) are available to I-SPY 2 TRIAL investigators. The frequency of patients reporting Adverse Events (AEs), alerts sent to providers, and the length of time for follow-up will be assessed in Phase 2 of the project.

Figure 3: Example plot and interactive interface provided to patients within the UCSF Breast Cancer Clinic setting showing an I-SPY 2 TRIAL patient PRO baseline result in relation to other I-SPY 2 TRIAL patients in the study.

At 18 I-SPY 2 TRIAL sites, three days prior to a specified visit (coinciding with a questionnaire), patients are notified and asked to complete the corresponding questionnaire. Participants who did not complete the questionnaire prior to their visit are provided with an iPad to complete the questionnaire in the waiting room and are supported by the front desk personnel. Once the scores are captured, the results are automatically processed and linked to other I-SPY 2 clinical trial data, normalized, and provided presented back to clinical investigators and the healthcare providers.

At the UCSF Breast Cancer Clinic sites, the ePRO solution has been designed for integration into the UCSF Epic EHR system, where the ePRO summary results will be presented back to I-SPY 2 TRIAL investigators through the EHR. This functionality was designed as part of the current project to assess feasibility and will go into production during Phase 2. Figure A2 in Appendix A shows the architecture for write back of I-SPY 2 TRIAL ePRO results to the UCSF site EHR that is in process for implementation.

DISCUSSION

Lessons Learned

The benefits of the OneSource project included the following:

- Better use of personnel
- Improved data quality through elimination of duplicate data entry and multiple data transformations
- Seamless integration of clinical care within a consolidated, secure framework

Many of the challenges encountered in this project can be overcome by increasing the value of electronic source data capture from electronic health records and other types of RWD. We plan to begin addressing these challenges in phase 2 of this project. A high-level overview of a future architecture and technology for OneSource is provided in the next section.

Any solution that can realistically be scaled to support EHR integration with EDC would also support the seamless, scalable integration with other clinical research information systems (e.g., ePRO, Clinical Trial Management System (CTMS) and Lab Information systems with an EDC) and the integration of the other systems with each other.

Future Work, Recommendations and Phase II Roadmap

In phase 2, key standards from CDISC and HL7 Standards Development Organizations (e.g., HL7 FHIR), controlled terminologies (e.g., Systematized Nomenclature of Medicine -- Clinical Terms (SNOMED CT), Logical Observation Identifiers Names and Codes (LOINC), Medical Dictionary for Regulatory Activities (MedDRA) and RxNORM) will be leveraged. In addition, National Cancer Institute (NCI) Cancer Data Standards Registry and Repository (caDSR) as well as its Enterprise Vocabulary Services (EVS) will be used.

In Phase II, OneSource will be extended to incorporate and integrate the following additional sources of data with the EHR (see Figure 4):

- a Substitutable Medical Applications, Reusable Technologies (SMART) on FHIR App launched within the UCSF EHR system to support source data capture of relevant data for the I-SPY 2 TRIAL (based on phase 1 progress)
- systems for the capture and visualization of patient reported outcome (PRO) data (PROMIS© and PRO-CTCAE™ for Quality of Life and adverse event reporting) that are integrated with the EHR for use at the point-of-care
- a mobile application for use by pathologists to capture pathological assessment data in structured format

The OpenClinica EDC, a validated 21 CFR Part 11 and Good Clinical Practice (GCP) compliant system, will be used for the capture of structured, standardized source data. The SMART App (Figure 4) will enable launch of the EDC system from within the EHR using single sign-on. The app will navigate to the correct patient within the EDC system via a secure, privacy-aware link between the EHR Medical Record Number and the EDC Study Participant ID for capture of relevant EHR.

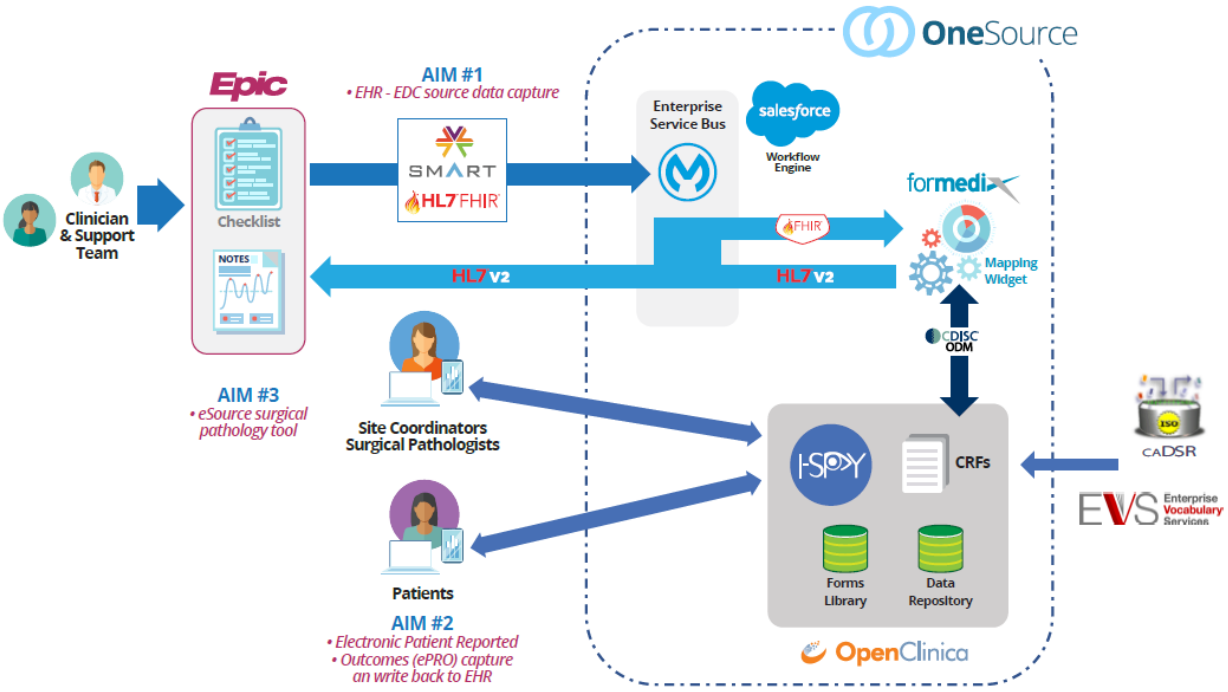


Figure 4: Architecture of proposed system addressing the project's three aims.

At UCSF, summary ePRO result data will be integrated and written back to the EHR system. In addition, an open source Adverse Event (AE) data explorer will be developed to analyze and view patient-reported adverse events compared to provider reported AEs.

The Formedix standards management platform will support rapid and efficient form generation and edit checks mappings and analysis results metadata using SDTM, ADaM, SEND, CDASH, ODM and NCI terminology. Formedix allows study design and management in an EDC platform-agnostic format that can rapidly generate forms in seven widely used EDC platforms, including OpenClinica, which is specifically used for the I-SPY 2 TRIAL. We will be able to generate forms in alternate EDC platform vendor formats as part of our deliverables.

Outcome measures for all components will assess efficiency in research use of clinical data; physician, clinical investigator, and patient satisfaction with ePRO system; data fidelity; and portability of software and data to other environments.

Future Collaboration

The OneSource team will collaborate with other eSource projects (e.g. Duke, UAMS, Memorial Sloan Kettering, biopharmaceutical industry, Society for Clinical Data Management (SCDM) and other stakeholders focused on using EHRs and other sources of RWD in clinical research.

CONCLUSION

Virtually every vision of the future of clinical research and care, whether it be from FDA (e.g., eSource Guidance, Use of EHR Data in Clinical Investigations Guidance and the Real World Evidence (RWE) Framework) [1,5,6], the NIH Roadmap, the Cancer Moonshot or the Learning Healthcare System, is predicated upon the ability to use high quality, accessible clinical data to accelerate the pace of research[1,5,6]. The challenge is enormous, given that clinical care and research remain separate silos of information, both in terms of connectivity and quality. While the majority of efforts to date have focused on technical integration to exchange data between EHRs and EDC systems, the emphasis must shift to the collection of high-quality data at the point-of-care. It is only when that issue is addressed that extraction of data in EHRs will provide the desired benefits.

The high quality of data achieved through use of EDC in the clinical trial environment has already established both the technological foundations and the template for implementation of this effort. Extending these same principles and technologies to the capture of data in clinical care will not only expedite improvements in clinical data quality, but also further the integration of cancer care and research by providing a common information infrastructure.

Integrating care and research will require a change in culture that begins at the point of care, where data are generated. The identification of key elements required for good and efficient care and the structuring of this data can bring great value to healthcare providers. We have worked to harmonize the data elements in early-stage breast cancer trials and clinical care and re-engineered our care processes to enable the efficient acquisition and display of data, supported by the tools that we have developed in this project. We have the technical tools and platforms for data capture and the mapping of that data to the appropriate standards for seamless secondary use. Using the UCSF Breast Care Center Program as the initial laboratory and the Athena Breast Health and WISDOM Network and the 20 site I-SPY 2 TRIAL to demonstrate extensibility and scalability, we have the ability to create a new path forward for achieving the vision of this program. Our goal is to demonstrate a more efficient, clinical system that integrates learning as a byproduct of care, while harmonizing data and enabling, through interoperability and consent processes, the efficient transfer of data to trials, in an analytic-ready format.

ACKNOWLEDGEMENTS

The OneSource team would like to thank Dr. Janet Woodcock, ShaAvhrée Buckman-Garner, Mitra Ahadpour, Mary Ann Slack, Mark Geanacopoulos, Wei Chen, Sara Meiselman, Tom Bechtold, Michael Ibara, Aheli Chattopadhyay, Alex Koorkoff, Karen Kimura, Jessie Hong, Garry Peterson, Victor Galvez, Matthew Shapiro, Heidi Collins, Smita Asare, Jeff Matthews, Becky Kush, Wayne Kubick, Gary Walker, Rhonda Facile and the I-SPY 2 TRIAL sponsor Quantum Leap Healthcare Collaborative, for their contributions to the project. In addition, we would like to thank Susan Lumsden and Scott Smith from HHS/ASPE for their support and guidance. In addition, we would like to thank the Salesforce and OpenClinica teams for their contribution to this project.

REFERENCES

1. Food and Drug Administration. Electronic Source Data in Clinical Investigations. fda.gov. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/electronic-source-data-clinical-investigations. Published September 2013. Accessed May 1, 2019.

2. 111th Congress of the United States of America. H.R.3590: Patient Protection and Affordable Care Act. congress.gov. https://www.congress.gov/bill/111th-congress/house-bill/3590. Published March 23, 2010. Accessed May 1, 2019.

3. Food and Drug Administration. Centers of Excellence in Regulatory Science and Innovation (CERSIs). fda.gov. https://www.fda.gov/science-research/advancing-regulatory-science/centers-excellence-regulatory-science-and-innovation-cersis. Published June 9, 2018. Accessed May 1, 2019.

4. CFR - Code of Federal Regulations Title 21 Part 11. fda.gov. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11. Accessed May 1, 2019.

5. Food and Drug Administration. Computerized Systems Used in Clinical Investigations: Guidance for Industry. fda.gov. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/computerized-systems-used-clinical-investigations. Published May 2017. Accessed May 2019.

6. Food and Drug Administration. Use of Electronic Health Record Data in Clinical Investigations. fda.gov. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry. Published July 2018. Accessed May 1, 2019.

7. US Department of Health and Human Services, Asst. Secretary of Planning and Evaluation. Patient-Centered Outcomes Research Trust Fund. aspe.hhs.gov. https://aspe.hhs.gov/patient-centered-outcomes-research-trust-fund. Published November 23, 2015. Accessed May 1, 2019.

8. Wrenn JO, Stein DM, Bakken S, Stetson PD. Quantifying clinical narrative redundancy in an electronic health record. *J Am Med Inform Assoc*. 2010;17(1):49-53. doi:10.1197/jamia.M3390.

9. Siegler EL, Adelman R. Copy and Paste: A Remediable Hazard of Electronic Health Records. *Am J Med*. 2009;122(6):495-496. doi:10.1016/j.amjmed.2009.02.010.

10. Kalra D, Schmidt A, Potts H, Dupont D, Sundgren M, De Moor G. Case Report from the EHR4CR Project—A European Survey on Electronic Health Records Systems for Clinical Research. *Health Connections*. 2011;1(2):108-113.

11. Retrieve Form for Data Capture - IHE Wiki. wiki.ihe.net. https://wiki.ihe.net/index.php/Retrieve_Form_for_Data_Capture. Published July 20, 2017. Accessed May 13, 2019.

APPENDICES

APPENDIX A: Technical Framework

The key architecture components in this figure include:

- **Epic Hyperspace** is the presentation component of the Epic suite. It is not a clinical module, but rather the actual application client that is presented to users of most areas of Epic. Clinical staff interact with Epic through Epic Hyperspace. When a healthcare provider or administrative staff launches Epic, the front-end software that is presented to them is called Hyperspace. For example, a physician will be presented with options to document clinical visits, place orders, and perform other clinically relevant tasks.
- **EPIC Interconnect for application integration**. Interconnect is a web service that exposes some Simple Object Access Protocol *(*SOAP) endpoints. SOAP is a standards-based Web services access protocol.
- **MuleSoft** is used by UCSF for integration of multiple applications including Epic EHR system. The goal is for MuleSoft to accelerate development of applications that are external to Epic by using a set of pre-built integration services. These templates save manual development work and leverage integration best practices. For the OneSource Pilot, MuleSoft provided support for various Epic interfaces as part of an Enterprise Service Bus architecture.
- **Salesforce** was originally the EDC platform for the I-SPY 2 TRIAL clinical trial platform. The team is now in the process of moving to OpenClinica for both the EDC and the ePRO platform based on the technology landscape assessment performed.

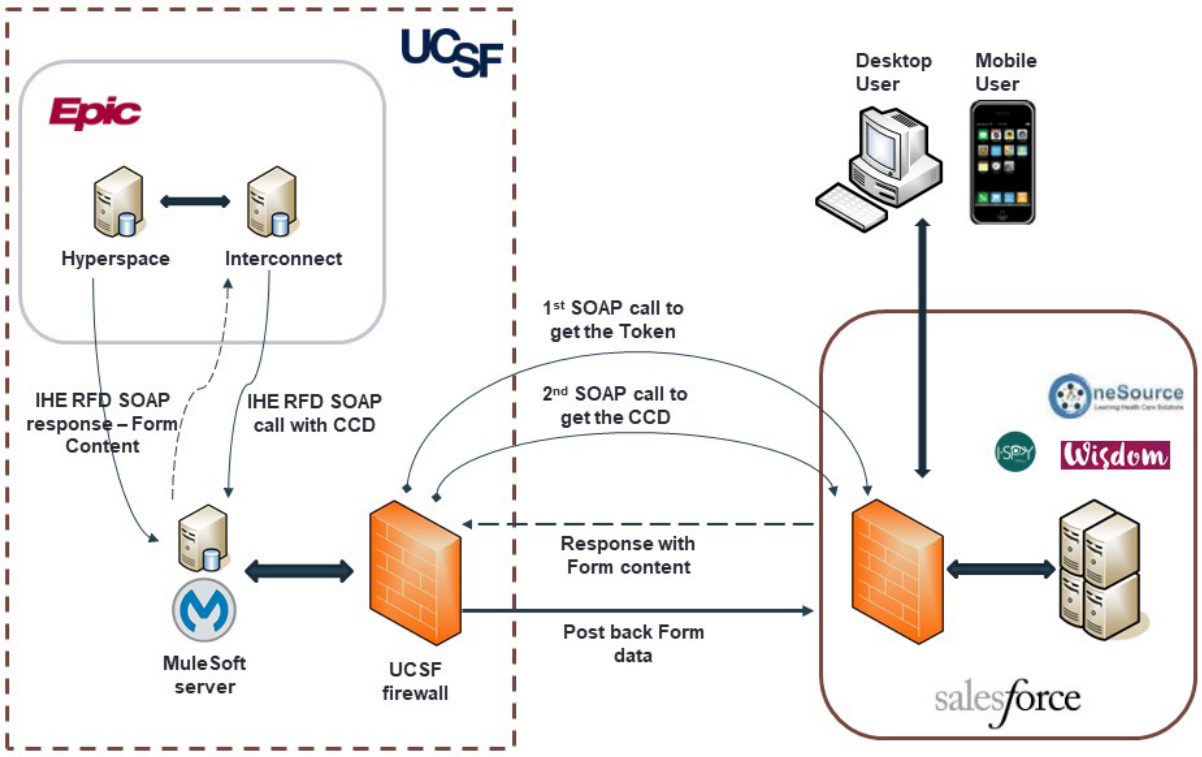


Figure A1: I-SPY 2 Data Exchange with Epic EHR Demonstration

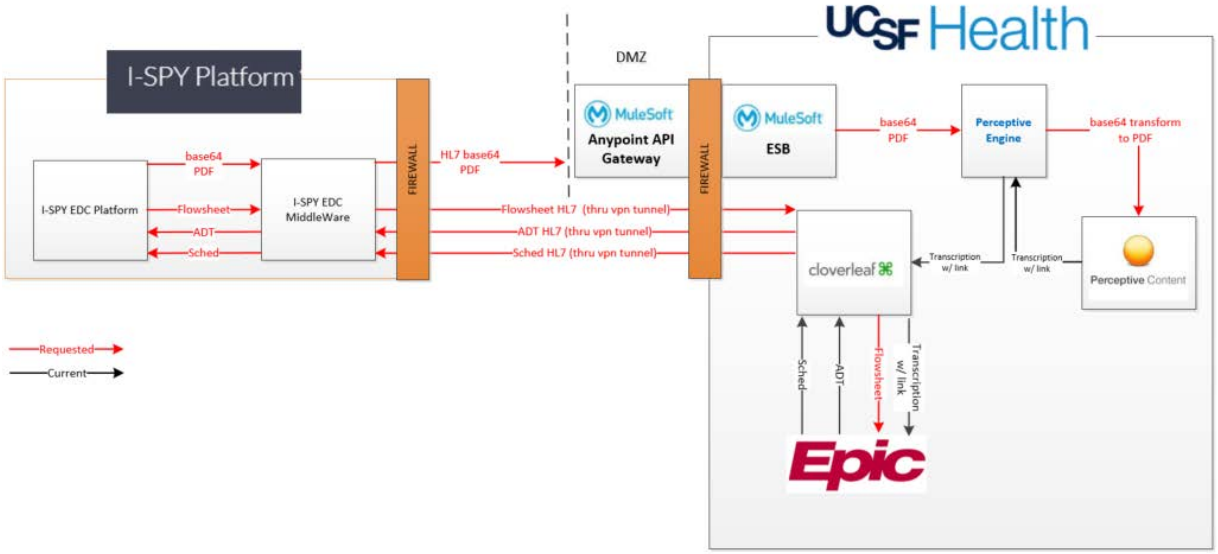


Figure A2: Integration of I-SPY 2 TRIAL platform for sending ePRO data back to the UCSF EHR system using standards.

APPENDIX B: Vendor Interview Questions and Selection Criteria

Questions used in the vendor evaluation and interview process were developed by I-SPY 2 TRIAL staff and contractors, and were broken down into several themes, as follows.

(a) Existing customer base

- How long has product been commercially available?
- What is the upgrade/new release cycle time?
- How many customers do you have on the platform?
- How does the study team interact with the platform? Web -based tool?
- Have customers submitted data to the FDA using your platform?
- Do investigators typically enter data directly or is data transcribed into the system from paper?
- Is the platform in use for commercial customers as well as academic medical centers?
- Are there any installations within the University of California system?

(b) Administration

- What is the typical deployment method? Cloud? local installation?
- Does the vendor provide services for deployment and administration?
- Does study set up require programming?
- Does the vendor provide support for study set up?
- Can a study be 'cloned' to initiate a new similar study?
- Does the platform make active use of existing standards such as CDISC, HL7 etc.?
- Do you have existing documentation on data security that could be provided to a security office?
- Is the platform 21 CFR part 11 compliant?
- Any experience working with validation teams?
- Do you provide full validation documents for the software?
- Does it use e-signatures? Have audit trails etc.?
- Can the platform be used in paper-less mode, where primary data entered directly to platform?
- Does your security model support different roles with unique access controls?

(c) Patient reported data

- Does the platform support direct interaction with patients for Patient Reported Outcome (PRO) and/or safety related findings?
- Does the platform include the more common PRO instruments like Quality of Life or does the study team have to create each PRO instrument?
- Can the PRO instrument be reused from one study to the next study?

(d) Reporting/Integration

- Any experience integrating clinical research with clinical practice EHR?
- Does the platform support structured clinical assessment, clinical summary from an EHR?

- Does the company have existing relationships with EPIC EHR or performed any integrations to that platform?
- Does the company have existing relationships with Salesforce or performed any integrations to that platform?
- Does the platform support integration with biomarkers and/or large genomic scale datasets?
- Is it possible to report/aggregate these data over more than one study?
- Do you have any integration with Google Cloud for analytics of data?
- Do you have customers using the platform as a precision medicine platform?

(e) Licensing and services

- What is the licensing model?
- One-time purchase, subscription, user based?
- If client installs and maintains locally, do you have documentation on infrastructure requirements?
- If the client installs and maintains locally, what staff would need to be dedicated to the platform to maintain it?
- How is support contracted?
- Do you provide audit support?
- Do you provide services for migrating data from existing platform to your platform?
- What is your Service Level Agreement time for responding to client support calls?

(f) Usability

- Is there access to presentations and/or videos of platform to review current use?
- If insufficient materials available online, is it possible to schedule a short demo?
- Evidence of 21 CFR part 11 compliance audit trails, electronic signatures?
- Evidence for PRO implementation?
- Is there evidence for direct interaction rather than paper transcribed to electronic?
- Is system readily useable by a study team? What is the look and feel? Incorporate experience with the platform here.

Assessment Results

(i) Company credentials:

Common: All vendors have both commercial and academic partners, some more of one than the other. The total number of customers is difficult to evaluate because some of the vendors have a small number of customers that are running many trials and other vendors have a larger number of customers running fewer trials each.

Distinguishing: The evaluation covered a wide of range of vendor experience from a high of 27 years to a low of just a few years.

(ii) EDC Architecture

Common: All vendors enable web-based data entry, preferably entered directly rather than transcribed from paper. All the vendors indicated that study set up did not require programming by the customer. All of the vendors provide:

- Study set up support
- Ability to copy one study to create a new one
- CDISC and/or HL7 standards
- Full set of security documents
- Audit trails and e-signatures
- Ability for system to be the source of information (paperless)

Distinguishing:

- Most of the vendors use Amazon Web Services (AWS) to host their cloud-based system.
- 21 CFR part 11 compliance is in place for most of the vendors

(iii) Patient Reported Outcomes

Common: Most of the vendors have incorporated ePRO directly into their EDC systems. All the vendors that provide ePRO enable the re-use of the instruments once defined,

Distinguishing: Two of the mature platforms have not implemented ePRO internally.

(iv) Reporting/Integration

Common: Most of the vendors state they can do integration with EHR systems at some level and support unstructured summaries. The vendors were not very familiar biomarker and genomic scale data.

Distinguishing: Experience with EHRs seems to be available from the EDC vendors.

A similar pattern emerges for reporting across studies. Most vendors responded that the client needs to export each study and merge them outside of the platform, or that it would require customization.

(v) Licensing and services

Common: All vendors have flexible licensing agreements from enterprise to per study arrangements and tiered pricing for academic and commercial. In addition, they all provide some level of support for migration from an existing EDC to a new one and can meet requirements of FDA audits and conform to FDA regulations.

Distinguishing: Service center response time depends on criticality of the issue and most of the vendors have SLA terms that specify times by criticality and appear reasonable.

(vi) Clinical Trial Management System (CTMS)

We conducted a research and captured CTMS functionality of the various vendors.

- Full suite of integrated or modular tools
- Reporting, monitoring, and query support
- Standalone CTMS
- Reporting only

APPENDIX C: IHE RFD SOAP Codebase Explanation

RFD specifies SOAP 12 web service transactions between a Form Filler and a Form Processor. The Filler requests a blank form from the Processor and submits the completed Case Report Form (CRF)to the Processor. In this integration, Epic acts as the Filler and OneSource as the Processor. The two RFD transactions that OneSource implemented are Retrieve Form and Submit Form. Each transaction consists of a request message and a response message. The content of these messages is depicted below.

RFD specifies SOAP 12 web service transactions between a Form Filler and a Form Processor. The Filler requests a blank form from the Processor and submits the completed Case Report Form (CRF) to the Processor. In this integration, Epic acts as the Filler and OneSource as the Processor. The two RFD transactions that OneSource implemented are Retrieve Form and Submit Form. Each transaction consists of a request message and a response message. The content of these messages as depicted below.

The OneSource Demonstration implemented a web service endpoint to respond to SOAP requests from Epic's RFD module. This endpoint serves the following Web Service Description Language (WSDL) message. The URL for this WSDL was configured into a button on an Epic screen. This URL was secure since SSL is supported by Epic, the OneSource web server, and the accompanying network infrastructure. The address **OneSourceSoapServer** was replaced with the URL of the OneSource SOAP endpoint. In addition to this WSDL, the OneSource web server provided an RFD.xsd schema file. This schema file was downloaded from the IHE FTP server and its use for guiding XML document generation is demonstrated in the XML message below.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <definitions xmlns="http://schemas.xmlsoap.org/wsdl/" xmlns:ihe="urn:ihe:iti:rfd:2007"
   xmlns:soap12="http://schemas.xmlsoap.org/wsdl/soap12/"
   xmlns:xsd="http://www.w3.org/2001/XMLSchema"
   xmlns:wsaw="http://www.w3.org/2005/08/addressing" name="FormProcessor"
   targetNamespace="urn:ihe:iti:rfd:2007">
3   <types>
4  <xsd:schema elementFormDefault="qualified" targetNamespace="urn:ihe:iti:rfd:2007">
5      <xsd:include schemaLocation="RFD.xsd"/>
6    </xsd:schema>
7   </types>
8
9   <message name="RetrieveForm_Message">
10   <part name="body" element="ihe:RetrieveFormRequest"/>
11  </message>
12  <message name="RetrieveFormResponse_Message">
13   <part name="body" element="ihe:RetrieveFormResponse"/>
14  </message>
15  <message name="SubmitForm_Message">
16   <part name="body" element="ihe:SubmitFormRequest"/>
17  </message>
18  <message name="SubmitFormResponse_Message">
19   <part name="body" element="ihe:SubmitFormResponse"/>
20  </message>
```

```
21
22  <portType name="FormProcessor_PortType">
23    <operation name="FormProcessor_RetrieveForm">
24      <documentation>Corresponds to Transaction ITI-34 of the IHE Technical
    Framework</documentation>
25      <input message="ihe:RetrieveForm_Message" wsaw:Action="urn:ihe:iti:2007:RetrieveForm"/>
26      <output message="ihe:RetrieveFormResponse_Message"
    wsaw:Action="urn:ihe:iti:2007:RetrieveFormResponse"/>
27    </operation>
28    <operation name="FormProcessor_SubmitForm">
29      <documentation>Corresponds to Transaction ITI-35 of the IHE Technical
    Framework</documentation>
30      <input message="ihe:SubmitForm_Message" wsaw:Action="urn:ihe:iti:2007:SubmitForm"/>
31      <output message="ihe:SubmitFormResponse_Message"
    wsaw:Action="urn:ihe:iti:2007:SubmitFormResponse"/>
32    </operation>
33  </portType>
34
35  <binding name="FormProcessor_Binding_Soap12" type="ihe:FormProcessor_PortType">
36    <soap12:binding style="document" transport="http://schemas.xmlsoap.org/soap/http"/>
37    <operation name="FormProcessor_RetrieveForm">
38      <soap12:operation soapAction="urn:ihe:iti:2007:RetrieveForm"/>
39      <input><soap12:body use="literal"/></input>
40      <output><soap12:body use="literal"/></output>
41    </operation>
42    <operation name="FormProcessor_SubmitForm">
43      <soap12:operation soapAction="urn:ihe:iti:2007:SubmitForm"/>
44      <input><soap12:body use="literal"/></input>
45      <output><soap12:body use="literal"/></output>
46    </operation>
47  </binding>
48
49  <service name="FormProcessor_Service">
50    <port binding="ihe:FormProcessor_Binding_Soap12" name="FormProcessor_Port_Soap12">
51      <soap12:address location="https://OneSourceSoapServer/RFDFormProcessor"/>
52    </port>
53  </service>
</definitions>
```

Retrieve Form Transaction

The IHE RFD Retrieve Form transaction (ITI-34) loads an RFD form from the Processor into the Filler.

Retrieve Form Request Message

The Filler (Epic) posts the following request message to the Processor (OneSource).

The ONESOURCE_CHECKLIST_ID is replaced with the identifier of the OneSource checklist. This string value is configured in the Epic button that the user presses to load the checklist. The ClinicalDocument element contains patient information in HL7 Clinical Document Architecture (CDA) format. This CDA

content is collected by Epic's RFD module. Patient identifiers are found in the XPath location //ClinicalDocument/recordTarget/patientRole/id.

```
<soap:Envelope xmlns:soap="http://www.w3.org/2003/05/soap-envelope">
 <soap:Header>
  <Action xmlns="http://www.w3.org/2005/08/addressing">urn:ihe:iti:2007:RetrieveForm</Action>
  <MessageID xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID</MessageID>
  <To xmlns="http://www.w3.org/2005/08/addressing">https://OneSourceSoapServer/RFDFormProcessor</To>
  <ReplyTo xmlns="http://www.w3.org/2005/08/addressing">
   <Address>http://www.w3.org/2005/08/addressing/anonymous</Address>
  </ReplyTo>
 </soap:Header>
 <soap:Body>
  <RetrieveFormRequest xmlns="urn:ihe:iti:rfd:2007">
   <prepopData>
    <ClinicalDocument>
     (Contains patient identifier)
    </ClinicalDocument>
   </prepopData>
   <workflowData>
    <formID>ONESOURCE_CHECKLIST_ID</formID>
    <encodedResponse>true</encodedResponse>
    <context/>
    <instanceID/>
   </workflowData>
  </RetrieveFormRequest>
 </soap:Body>
</soap:Envelope>
```

Retrieve Form Response Message

The Processor (OneSource) responds to the above request by sending the HTML checklist to the Filler (Epic) within the message below.

The **Structured element** contains the HTML/CSS markup to render the OneSource checklist in Epic. The HTML must conform to XHTML Basic and W3C HTML Compatibility Guidelines provided in the Appendix C of the W3C XHTML 1.0 Recommendation. The Epic RFD module's embedded web browser might impose further constraints on the markup.

The **ONESOURCE_CHECKLIST_INSTANCE_ID** specifies the unique instance of the checklist for the requested patient.

```
<soap:Envelope xmlns:soap="http://www.w3.org/2003/05/soap-envelope">
 <soap:Header>
```

```
    <Action xmlns="http://www.w3.org/2005/08/addressing">urn:ihe:iti:2007:RetrieveFormResponse</Action>
    <MessageID xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID1</MessageID>
    <To xmlns="http://www.w3.org/2005/08/addressing">http://www.w3.org/2005/08/addressing/anonymous</To>
    <RelatesTo xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID2</RelatesTo>
  </soap:Header>
  <soap:Body>
    <RetrieveFormResponse xmlns="urn:ihe:iti:rfd:2007">
      <form>
        <Structured>
          <html><head>(JavaScript)</head><body>(Checklist and Submit button)</body></html>
        </Structured>
        <instanceID>ONESOURCE_CHECKLIST_INSTANCE_ID</instanceID>
      </form>
      <contentType>application/xhtml+xml</contentType>
      <responseCode>SUCCESS</responseCode>
    </RetrieveFormResponse>
  </soap:Body>
</soap:Envelope>
```

Retrieve Form Error Message

The Processor responds to insufficient requests (such as unidentified patient) with the following SOAP error message. If the requested **ONESOURCE_CHECKLIST_ID** is invalid, the reason should be Unknown formID.

```
<env:Envelope xmlns:env=http://www.w3.org/2003/05/soap-envelope xmlns:xml="http://www.w3.org/XML/1998/namespace">
  <env:Body>
    <env:Fault>
      <env:Code>
        <env:Value>env:Sender</env:Value>
      </env:Code>
      <env:Reason>
        <env:Text xml:lang="en">Required Information Missing</env:Text>
      </env:Reason>
    </env:Fault>
  </env:Body>
</env:Envelope>
```

Skipping Checklist Questions

The checklist requires skip logic to hide questions that are contingent upon previous answers. Implementation of this skip logic depends upon the JavaScript capabilities of the Epic RFD embedded browser component. JavaScript might be employed to implement skip logic by setting display properties. Alternatively, the checklist may be divided into a series of forms, each on its own web page.

In this case, the skip logic determines the sequence and is executed by the Processor when each form is submitted.

Submit Form

The IHE RFD Submit Form transaction (ITI-35) submits data collected by the Filler to the Processor.

Submit Form Request Message

The HTML form loaded into the Filler (Epic) will post the following SOAP request message to the Processor (OneSource).

The SubmitFormRequest element will contain the **ONESOURCE_CHECKLIST_INSTANCE_ID** and the answers in XML format. The syntax of this element is determined by OneSource's requirements.

```
<soap:Envelope xmlns:soap="http://www.w3.org/2003/05/soap-envelope">
 <soap:Header>
  <Action xmlns="http://www.w3.org/2005/08/addressing">urn:ihe:iti:2007:SubmitForm</Action>
  <MessageID xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID</MessageID>
  <To xmlns="http://www.w3.org/2005/08/addressing">https://OneSourceSoapServer/RFDFormProcessor</To>
  <ReplyTo xmlns="http://www.w3.org/2005/08/addressing">
   <Address>http://www.w3.org/2005/08/addressing/anonymous</Address>
  </ReplyTo>
 </soap:Header>
 <soap:Body>
  <SubmitFormRequest xmlns="urn:ihe:iti:rfd:2007">
   <OneSourceChecklist id="ONESOURCE_CHECKLIST_INSTANCE_ID">(checklist answers)</OneSourceChecklist>
  </SubmitFormRequest>
 </soap:Body>
</soap:Envelope>
```

Submit Form Response Message

The Processor shall return the HTTP response code 200 – OK to indicate success. If the Processor cannot recognize the posted data, then the Processor shall return the HTTP response code 400 – Bad Request. The Filler displays the content element of the response from the Processor. These results may be either a report or a subsequent HTML form. The latter is used if the checklist is divided into a sequence of HTML forms whose skip logic is implemented on the Processor after each form is submitted. In either case, the HTML must conform to XHTML Basic and W3C HTML Compatibility Guidelines provided in the Appendix C of the W3C XHTML 1.0 Recommendation. The Epic RFD module's embedded web browser might impose further constraints on the markup.

```
<soap:Envelope xmlns:soap="http://www.w3.org/2003/05/soap-envelope">
 <soap:Header>
```

```
30    <Action
xmlns="http://www.w3.org/2005/08/addressing">urn:ihe:iti:2007:SubmitFormResponse</Action>
31    <MessageID xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID1</MessageID>
32    <To
xmlns="http://www.w3.org/2005/08/addressing">http://www.w3.org/2005/08/addressing/anony
mous</To>
33    <RelatesTo xmlns="http://www.w3.org/2005/08/addressing">urn:uuid:GUID2</RelatesTo>
34   </soap:Header>
35   <soap:Body>
36    <SubmitFormResponse xmlns="urn:ihe:iti:rfd:2007">
37     <content>
38      <Structured>
39       <html><head>(JavaScript)</head><body>(Report or subsequent form)</body></html>
40      </Structured>
41      <instanceID>ONESOURCE_CHECKLIST_INSTANCE_ID</instanceID>
42     </content>
43     <contentType>application/xhtml+xml</contentType>
44     <responseCode>OK</responseCode>
45    </SubmitFormResponse>
46   </soap:Body>
47 </soap:Envelope>
```

APPENDIX D: CRF/ePRO IHE RFD Codebase Resources

I-SPY 2 TRIAL ePRO forms
PROMIS and PRO-CTCAE questionnaires in x-form format
https://mobile.athenacarenetwork.org/OneSource/Archive.zip

I-SPY 2 TRIAL Case Report Forms (CRFs)
Lab & Test, Baseline Symptoms and CRFs forms used in assessment
https://mobile.athenacarenetwork.org/OneSource/ISPY-CRFs.zip

IHE-RFD SOAP Interconnect codebase
Code used of IHE RFD Epic connection
https://mobile.athenacarenetwork.org/OneSource/ucsf-iherfd.zip

APPENDIX E: Detailed Gap Analysis between Case Report Forms (CRFs) and UCSF EHR system

Table E1: I-SPY 2 CRF Data Elements for Lab and Test Results and UCSF EHR system (EPIC)

Field Label	Found in Epic?	Epic Production Location(s)	Easy to find in Epic	Structured?	Comments
Leukocytes	Y	Labs > CBC	Y	?	WBC count; Need to verify structure
Absolute Neutrophils	Y	Labs > Differential	Y	Y	
Platelets	Y	Labs > CBC	Y	Y	
Total Bilirubin	Y	Labs > Liver profile	Y	Y	
AST	Y	Labs > Liver profile	Y	Y	ASpartate Transaminase
ALT	Y	Labs > Liver profile	Y	Y	ALanine Transaminase
Creatinine	Y	Labs > Chem profile	Y	Y	
MUGA Scan	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo"	N	N/A	Either MUGA Scan or Cardiac Echo may be present as they are the "same test"; users fill out one or the other using the "TransThoracic Echo" result
MUGA Scan Date of Procedure	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo"	N	N/A	
MUGA Scan LVEF %	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo" > Conclusions Section	N	N/A	
MUGA Scan LVEF % Institution lower limit of normal	inconsistently entered		N	N/A	
Cardiac Echo	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo"	N	Y	
Cardiac Echo Date of Procedure	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo"	N	Y	in label
Cardiac Echo LVEF%	inconsistently entered	Chart Review > Cardiology > "TransThoracic Echo" > Conclusions Section	N	Y	
Cardiac Echo LVEF % Institution lower limit of normal	inconsistently entered		N	?	

Table E2: Gap Analysis between the I-SPY 2 Lab and Test Results CRF and UCSF EHR system (EPIC)

I-SPY 2 TRIAL Case Report Form Data Element	CDISC CDASH	Structured Data in EPIC
Leukocytes	LBORRES / LBORRESU where LBTESTCD="WBC", LBTEST="Leukocytes" and LBORRESU="10^9/L"	Y
Absolute Neutrophils	LBORRES / LBORRESU where LBTESTCD="NEUT", LBTEST="Neutrophils" and LBORRESU="10^9/L"	Y
Platelets	LBORRES / LBORRESU where LBTESTCD="PLAT", LBTEST="Platelets" and LBORRESU="10^9/L"	Y
Total Bilirubin	LBORRES / LBORRESU where LBTESTCD="BILI", LBTEST="Bilirubin" and LBORRESU="mg/dL"	Y
Aspartate Aminotransferase (AST)	LBORRES / LBORRESU where LBTESTCD="AST", LBTEST="Aspartate Aminotransferase" and LBORRESU="U/L"	Y
Alanine Aminotransferase (ALT)	LBORRES / LBORRESU where LBTESTCD="ALT", LBTEST="Alanine Aminotransferase" and LBORRESU="U/L"	Y
Creatinine	LBORRES / LBORRESU where LBTESTCD="CREAT", LBTEST="Creatinine" and LBORRESU="mg/dL"	Y
Multigated Acquisition (MUGA) Scan	CVMETHOD="MUGA"	N/A
MUGA Scan Date of Procedure	CVDTC	N/A
MUGA Scan LVEF %	CVORRES / CVORRESU where CVTESTCD="LVEF", CVTEST="Left	N/A

	Ventricular Ejection Fraction" and CVORRESU="%"	
MUGA Scan LVEF % Institution lower limit of normal	CVORNRHI and CVORNRLO where CVTESTCD="LVEF", CVTEST="Left Ventricular Ejection Fraction" and CVORRESU="%"	N/A
Cardiac Echo (echocardiogram)	CVMETHOD= "ECHOCARDIOGRAPHY"	Y
Cardiac Echo Date of Procedure	CVDTC	Y
Cardiac Echo LVEF%	CVORRES / CVORRESU where CVTESTCD="LVEF", CVTEST="Left Ventricular Ejection Fraction" and CVORRESU="%"	Y
Cardiac Echo LVEF % Institution lower limit of normal	CVORNRHI and CVORNRLO where CVTESTCD="LVEF", CVTEST="Left Ventricular Ejection Fraction" and CVORRESU="%"	N/A
Total		**10/15 (67%)**

Table E3: Annotated I-SPY 2 CRF for Labs and Test Results with CDISC CDASH variables

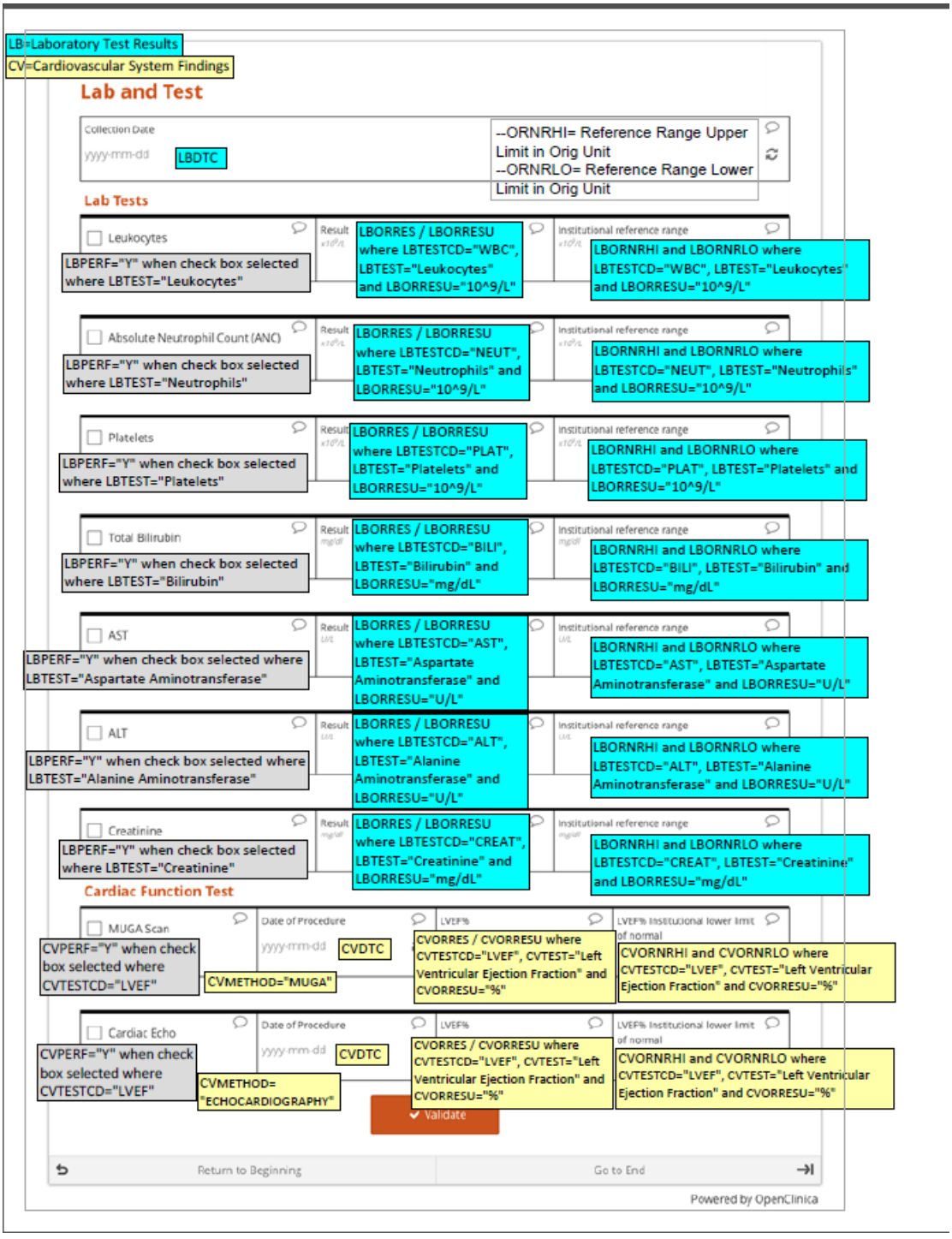


Table E4: I-SPY 2 CRF Data Elements for Menopausal Status and UCSF EHR system (EPIC)

I-SPY 2 Field	Found in Epic	Epic Production Location(s)	Easy to find in Epic	Structured?	Comments
Date of last menstrual period	Inconsistently Entered	History > Social Documentation (free text)	Y	N	Maybe in patient age form
On estrogen replacement	Inconsistently Entered	History > Social Documentation (free text)	N	N	Might be found in medications, if updated properly
Duration of estrogen replacement	Inconsistently Entered	History > Social Documentation (free text)	N	N	Might be found in medications, if updated properly
Bilateral Oophorectomy	Inconsistently Entered	History > /Surgical	N	N	Must look in surgery, if any, but could be elsewhere
Date of Bilateral Oophorectomy	Inconsistently Entered	History > /Surgical	N	N	Must look in surgery, if any, but could be elsewhere
Hysterectomy	Inconsistently Entered	History > /Surgical	N	Y	Must look in surgery, if any, but could be elsewhere
Date of Hysterectomy	Inconsistently Entered	History > /Surgical	N	Y	Must look in surgery, if any, but could be elsewhere
Menopausal Status	Inconsistently Entered	History > Social Documentation (free text)	N	N	Maybe in patient age form

Table E5: Gap Analysis between the I-SPY 2 Menopausal Status CRF and UCSF EHR system (EPIC)

I-SPY 2 TRIAL Case Report Form Data Element	CDISC CDASH Standard	Structured Data in EPIC
Date Last Menstrual Period	RPORRES="UNKNOWN DATE" when RPTESTCD="LMPSTKNW" and RPTESTCD="Last Menstrual Period Start Date Known"	N
On Estrogen Replacement	CMDUR / CMDURU	N
Estrogen Replacement Duration Months	CMDUR / CMDURU	N
Estrogen Replacement Duration Years	CMDUR2 / CMDURU2	N
Bilateral oophorectomy	PROCCUR="Y" when PRTRT="BILATERAL OOPHERECTOMY"	N
Bilateral Oophorectomy Date	PRSTDAT	N
Hysterectomy	PROCCUR="Y" when PRTRT="HYSTERECTOMY"	Y
Hysterectomy Date	PRSTDAT	Y
Menopausal Status	RPORRES="UNKNOWN DATE" when RPTESTCD="LMPSTKNW" and RPTESTCD="Last Menstrual Period Start Date Known"	N
On Estrogen Replacement	CMOCCUR="Y" when CMTRT="ESTROGEN REPLACMENT"	*
Total Matching		**2/11 (18%)**

* (depends on if under Medications or Social Documentation)

Table E6: Annotated I-SPY 2 CRF for Menopausal Status with CDISC CDASH variables

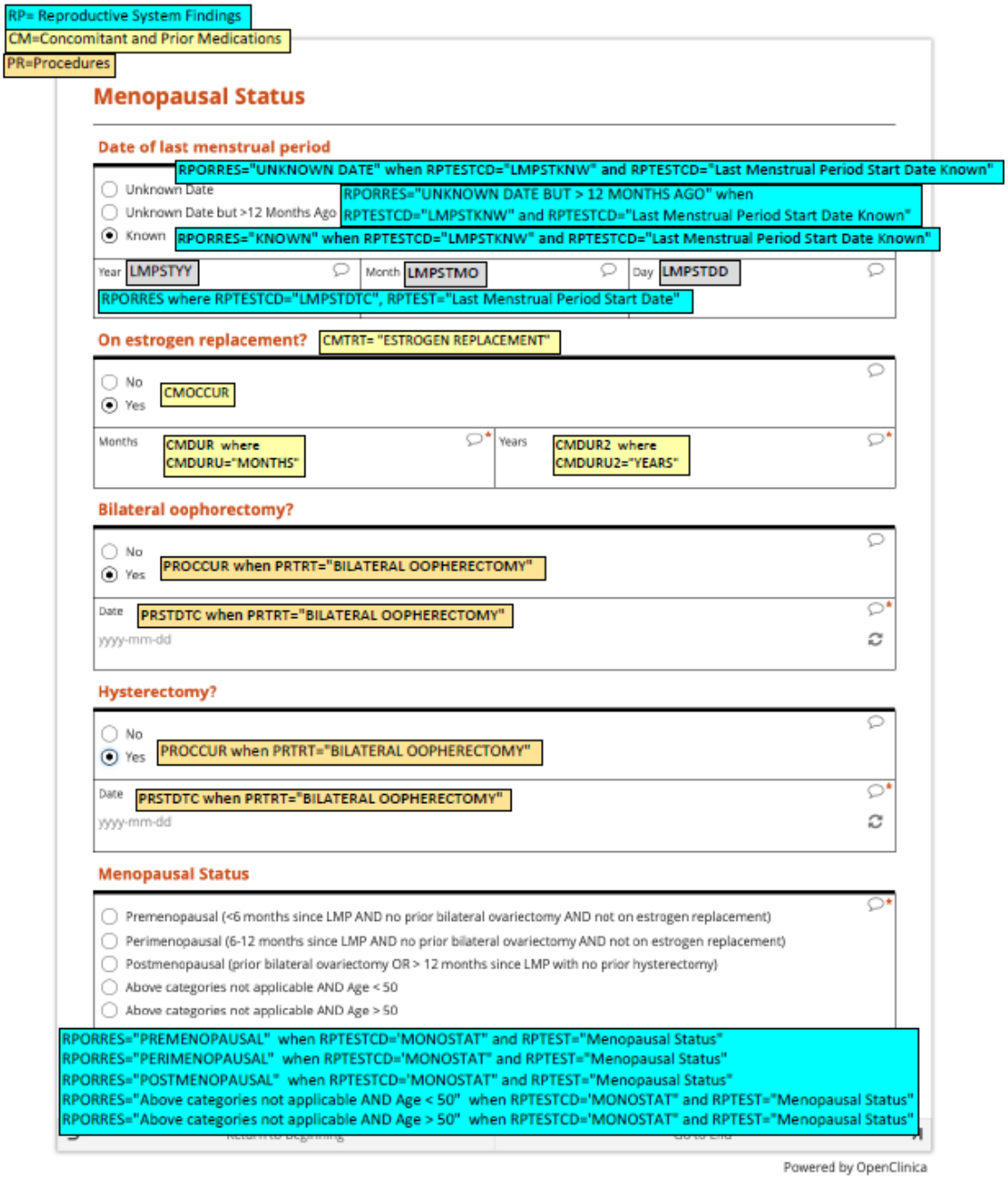


Table E7: I-SPY 2 CRF Data Elements for Baseline Symptoms and UCSF EHR system (EPIC)

I-SPY 2 Field	Found in Epic?	Epic Production Location(s)	Easy to Find in Epic?	Structured?	Comment
Repeating Group: Allergies					
Allergy Type	N	Not present	N/A	N/A	Reaction Type may be alternative
Allergy Name	Y	Allergies/Contraindications > Agent	N	Y	
Reaction(s)	Inconsistently entered	Allergies/Contraindications > Reactions	N	Y	May be in allergies section or in note
Severity	Inconsistently entered	Allergies/Contraindications > Severity	N	N/A	Frequently not entered directly in clinic note
Year first seen	Inconsistently entered	Not present (Date Noted Present)	N	N/A	Frequently not entered directly in clinic note
Baseline Condition					
Type	N	Not present	N/A	N/A	
Condition	Y	History > Medical History > Problem (coded)	N	Y	
Severity Grade	Inconsistently entered	Not present	N	N/A	Frequently not entered directly in clinic note or history
Attribution	Inconsistently entered	N/A (should always be baseline)	N	N/A	Frequently not entered directly in clinic note or history
Onset Date MM	Inconsistently entered	History > Medical History > Date	N	Y	
Onset Date DD	Inconsistently entered	History > Medical History > Date	N	Y	
Onset Date YYYY	Inconsistently entered	History > Medical History > Date	N	Y	
Resolved	Inconsistently entered	Not present	N	N	
End Date MM	N	Not present	N	N/A	
End Date DD	N	Not present	N	N/A	
End Date YYYY	N	Not present	N	N/A	
Baseline Symptoms					
No label, but select symptom type	Inconsistently entered	Clinical progress note	Y or N	N	Categorized by type of symptom Y only if abnormal labs
Grade	N	Not present	N	N/A	Not in Epic; graded by coordinator based on CTCAE
Attribute	N	N/A	N	N/A	Not in Epic; entered in EDC by coordinator working with physician
Was event life threatening at time of event?	N	Clinical progress note	N	N/A	Not in Epic; usually not in note unless hospitalized
Did event require inpatient hospitalization or prologation of existing hospitalization?	N	Clinical progress note	Y	N	
Did event result in persistent or significant disability/incapacity or substantial disruption of the ability to perform life functions?	Inconsistently entered	Clinical progress note	Y or N	N	Only if hospitalized for AE
Did event result in congenital abnormality/birth defect?	Inconsistently entered	Clinical progress note	Y or N	N	Not found unless hospitalized
Did the investigator find this event very unusual and/or potentially serious, but didn't meet any of the above criteria?	Inconsistently entered	Clinical progress note	Y or N	N	Not found unless hospitalized
Did event result in death?	Inconsistently entered	N/A	Y or N	N/A	Not found unless hospitalized
AE Onset Date	Inconsistently entered	Date on clinical progress note	Y or N	N	Sometimes difficult to find; sometimes missing altogether
Resolved	Inconsistently entered	TBD; might be at end of clinic progress note, per Lauren Dickey	Y or N	N	May not be in clinic note; multiple places for AE entry
AE End date	Inconsistently entered	TBD; might be at end of clinic progress note, per Lauren Dickey	Y or N	N	Sometimes difficult to find; sometimes missing altogether
Adverse Event Special Interest	Inconsistently entered	Not present	N	N/A	
Immune Related Adverse Event	Y	Clinical progress note	Y	N	
Patient Complaint	N	Clinical progress note	N	N	

Table E8: Summary of mappings for data elements in Baseline CRF

Group	Allergies	Baseline Conditions	Baseline Symptoms
Total # of Data Elements	5	11	15
# of Data Elements where Epic maps 1:1 with I-SPY 2 TRIAL	2 Data Elements • Allergy Name • Allergic Reaction	4 Data Elements • Condition • Onset MM • Onset DD • Onset YY	Zero (0) Data Elements
# of Data Elements not recorded at all in Epic	1 Data Element • Allergy Type	4 Data Elements • Baseline Type • End Date MM • End Date DD • End Date YY	2 Data Elements • Grade • Patient Compliant
# of Data Elements inconsistently defined in epic	3 Data Elements • Reactions • Severity • Year First Seen	5 Data Elements • Severity Grade • Attribution • Onset MM • Onset DD • Onset YY	11 Data Elements • Did the event result in disability/incapacity…? • Did the event result in congenital abnormality/birth defect? • Did the investigator find this event very unusual… serious…? • Did event result in death? • Onset Date • Resolved • End Date • Adverse Event Special Interest • Attribute • Was the event life threatening at the time of event? • Did event require in-patient hospitalization?
# of Data Elements difficult to find in epic.	5of 5 (100%)	11of 11 (100%)	• 13 of 15 (87%) difficult to find in Epic except • Did event require in-patient hospitalization? • Immune-related Adverse Event

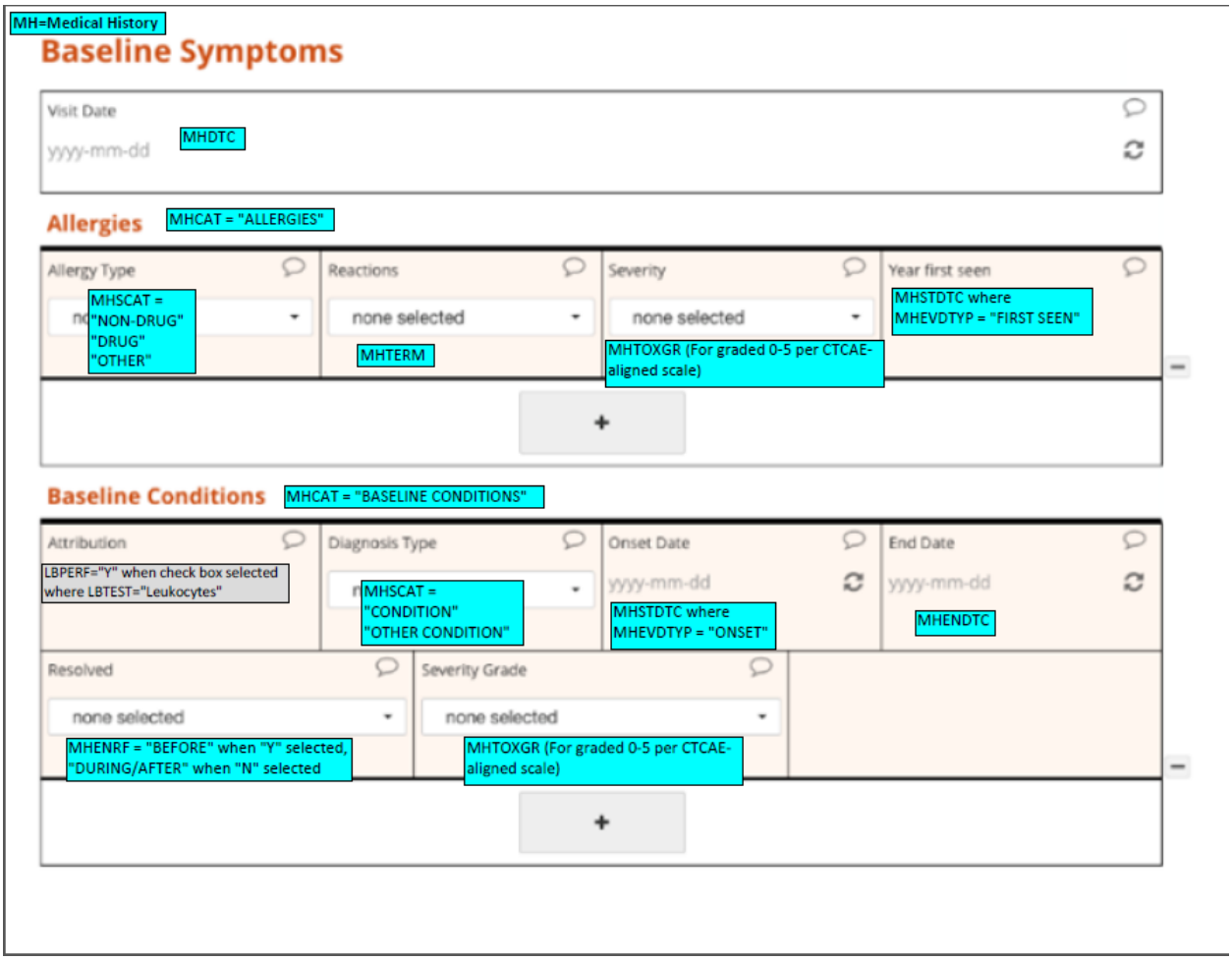
Table E9: Gap Analysis between the I-SPY 2 Baseline Symptoms CRF and UCSF EHR system (EPIC)

I-SPY 2 TRIAL Case Report Form Data Element	CDISC CDASH Standard	Data in EPIC
Repeating Group: Allergies	**Medical History and Adverse Event Domains**	
Allergy Type	MHCAT	N/A
Allergy Name	MHTERM	Y
Reaction(s)	MHSCAT	Y
Severity	MHTOXGR	Y
Year first seen	MHSTDTC where MHEVDTYP = "FIRST SEEN"	N/A
Baseline Condition	**Medical History (MH) Domain**	
Type	MHCAT	N/A
Condition	MHSCAT	Y
Severity Grade	MHTOXGR	N/A
Attribution	Not present	N/A
Onset Date MM	MHSTMO where MHEVDTYP = "ONSET"	Y
Onset Date DD	MHSTDD where MHEVDTYP = "ONSET"	Y
Onset Date YYYY	MHSTYY where MHEVDTYP = "ONSET"	Y
Resolved	MHONGO = “No”	N/A
End Date MM	MHENMO	N/A
End Date DD	MHENDD	N/A
End Date YYYY	MHENYY	N/A
Baseline Symptoms (categorized by type of symptom)	**Adverse Event (AE) Domain**	
No label, but select symptom type		N*
Grade	AETOXGR (Tox. Grade)	N
Attribute		N/A
Was event life threatening at time of event?	AESLIFE	N
Did event require inpatient hospitalization or prolongation of existing hospitalization?	AESHOSP	N
Did event result in persistent or significant disability/incapacity or substantial disruption of the ability to perform life functions?	AESDISAB	N
Did event result in congenital abnormality/birth defect?	AESCONG	N
Did the investigator find this event very unusual and/or potentially serious, but didn't meet any of the above criteria?	AESER	N
Did event result in death?	AESDTH	N/A
AE Onset Date	AESTDAT, and AESTTIM	N
Resolved	AEONGO = “N"	N

AE End date	AEENDAT and AEENTIM	N
Adverse Event Special Interest	Not Present	N/A
Immune Related Adverse Event	Not Present	N
Patient Complaint	Not Present	N
Total Matching		**7/30 (22%)**

* (but Y if abnormal labs are generated from the flowsheet)

Table E10: Annotated I-SPY 2 CRF for Baseline Symptoms with CDISC CDASH variables



APPENDIX F: Electronic Patient Reported Outcome (ePRO)Survey Questions

Questionnaires	Source
Global Measures (3 questions: QoL, fatigue, pain)	Mayo clinic
FACT-G (1 question treatment)	FACT-G
Distress Thermometer	NCCN
Fear of Recurrence	NCCN
Physical Function	PROMIS
Anxiety	PROMIS
Depression	PROMIS
Fatigue	PROMIS
Cognitive Function	PROMIS
Social Roles	PROMIS
Sexual Interest/ Function	PROMIS
Pain Interference	PROMIS
Sleep Disturbance	PROMIS
PROPr (utility score)	PROMIS
Full-Set	PRO-CTCAE
Dry eye questions	Focus groups

APPENDIX G: CDASH Data Elements Definitions

I-SPY 2 Lab and Test Results CFR

CDASH Variable	CDASH Variable Label
LBORRES	Lab Result
LBORRESU	Lab Result Unit
LBTESTCD	Lab Test or Examination Short Name (SDTM variable used as needed by CDASH)
LBTEST	Lab Test or Examination Name
CVMETHOD	Method of the test or examination.
CVDTC	Date/Time of Test (SDTM variable used as needed by CDASH)
CVORRES	Result or Finding in Original Units
CVORRESU	Original Units
CVTESTCD	Short Name of Cardiovascular Test (SDTM variable used as needed by CDASH)
CVTEST	Name of Cardiovascular Test
CVORNRHI	Normal Range Upper Limit- Original Unit
CVORNRLO	Normal Range Lower Limit- Original Unit

I-SPY 2 Menopausal Status CRF

CDASH Variable	CDASH Variable Label
RPORRES	Reproductive System Findings Result
RPTESTCD	Short Name of Reproductive Test (SDTM variable used as needed by CDASH)
CMCDUR	Collected Duration (*NOTE: the CDASH variable includes the “c” for “collected”)
CMCDURU	Collected Duration Unit (*NOTE: the CDASH variable includes the “c” for “collected”)
PROCCUR	Procedure Occurrence
PRTRT	Procedure Name
PRSTDAT	Procedure Start Date
CMOCCUR	Concomitant Meds Occurrence
CMTRT	Concomitant Medication Name

I-SPY 2 Baseline CRF

CDASH Variable	CDASH Variable Label
RPORRES	Reproductive System Findings Result
RPTESTCD	Short Name of Reproductive Test (SDTM variable used as needed by CDASH)
CMDUR	Collected Duration (*NOTE: the CDASH variable includes the “c” for “collected”)
CMDURU	Collected Duration Unit (*NOTE: the CDASH variable includes the “c” for “collected”)
PROCCUR	Procedure Occurrence
PRTRT	Procedure Name
PRSTDAT	Procedure Start Date
CMOCCUR	Concomitant Meds Occurrence
CMTRT	Concomitant Medication Name