AWARD NUMBER: W81XWH-14-1-0111

TITLE: A Molecular Framework for Understanding DCIS

PRINCIPAL INVESTIGATOR: Herbert Lyerly

CONTRACTING ORGANIZATION: Duke University
Durham, NC 27705

REPORT DATE: October 2016

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
October 2016	Annual	30 Sep 2015 - 29 Sep 2016

A Molecular Framework for Understanding DCIS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-14-1-0111

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
H. Kim Lyerly, M.D.

E-Mail: kim.lyerly@dm.duke.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Duke University
2200 W. Main St., STE 700
Durham, NC 27705

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

DCIS is proposed to be a precursor to invasive breast cancer. Improvements in early diagnosis have led to increased numbers of DCIS cases, however, we presently have no way to predict which DCIS lesions are at risk for progression to invasive cancer. We also lack the extensive molecular profiling necessary to place DCIS within the framework used to classify and guide treatment for invasive disease. We propose to take advantage of a unique set of specimens, comprising DCIS and early invasive disease, and using next-generation sequencing, understand the nature of DCIS and the events that determine and promote its progression. We have already developed the approaches necessary for obtaining profiles from the tumor and stromal compartments of DCIS, lesions from which only limited numbers of cells can be obtained. The availability of our datasets will transform the understanding of early disease and may ultimately alter the course of treatment for women with a diagnosis of DCIS.

15. SUBJECT TERMS

Breast cancer, DCIS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
U					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified	28	

Table of Contents

Page

A Molecular Framework for Understanding DCIS

Award No. W81XWH-14-1-0111
Annual Report Year 2

1. Introduction

This project centers on creating a molecular framework of DCIS (ductal carcinoma in situ). DCIS is considered to be the precursor to Invasive Ductal Carcinoma (IDC), the most common form of breast cancer. IDC accounts for 80% of all breast cancers, predominantly affecting women aged 55 and older; however, at least a third of women with IDC are diagnosed before they reach 55. Utilizing a unique bank of frozen mammary biopsies, containing samples with DCIS alone, and a combination of DCIS and IDC, we aim to profile both DCIS and related tissue components. It is our aim to sample the ~300 biopsies, and compare both by RNA seq, and whole genome amplification, DCIS lesions, within, and between patients, and see how these may be correlated with IDC lesions. We also intend to look for changes in the stroma between those patients that present with IDC and those that do not. This work aims to identify characteristics that may be suggestive of a patients' likelihood of progressing from DCIS to IDC, with the purpose of reducing the need for over treatment for this disease.

2. Approach

We first applied and received appropriate regulatory approval from the Duke SPORE tissue use committee, the Duke Cancer Center Protocol Review Committee, the Duke IRB, and the DoD. An MTA was established with the University of Cambridge.

We have created a DCIS Clinical Annotation database which was created and is being managed by Cedars-Sinai Medical Center in Los Angeles. A remote, web based case report from was also created to enable data abstraction for Duke medical records.

All Duke breast cancer research specimens were identified and prioritized based on pathologic diagnosis. Clinical data was abstracted from Duke clinical data by prioritizing pure DCIS, mixed DCIS and invasive BC samples, and invasive BC samples. Normal breast and atypical will be annotated if needed as controls.

Data was gathered from 1) electronic medical records at Duke University Medical Center and 2) extracted from DCIS patient charts (which hold the specific clinical annotation including patient outcome).

After identifying breast cancer research specimens with pathologic and clinical data consistent with DCIS and/or DCIS and invasive breast cancer, shipping to Dr. Greg Hannon for further processing was performed.

3. Key Research Accomplishments

Major goals of the project (as stated in SOW)

1. Regulatory approval and MTA (Duke/Cambridge)
2. Database creation and remote web based CRF created (Duke)
3. Sample collection/annotation, shipping to Cambridge (Duke)
4. Laser capture of frozen material (Duke pathologist working remotely or at CSHL/Cambridge)
5. Exome capture and DNA sequencing (CSHL/Cambridge)
6. RNAseq library construction (CSHL/Cambridge)
7. Analysis DNA data (CSHL/NYGC)
8. Analyze RNA differential expression (NYGC)
9. Analyze stroma compartments (CSHL/Cambridge)
10. Technical validation of potential markers (Cambridge)
11. Validate potential markers in FFPE cohort (Duke/Cambridge)
12. Validate in longitudinal cohort (Duke /Cambridge)
13. Nominate candidates for clinical validation (Duke /Cambridge)

Dr. Greg Hannon of Cancer Research UK Cambridge Institute will provide his research accomplishments via a separate report.

What was accomplished under these goals:

Tissue Use protocol

A Tissue Use protocol, "Molecular Framework of Early Breast Cancer" (IRB# Pro00059726) was approved by the Duke IRB, Cancer Protocol Committee (CPC), and DOD to allow for the use of banked tissue samples for the proposed studies (see Appendix for protocol). The 2016 continuing review by the Duke IRB is included in the Appendix as well.

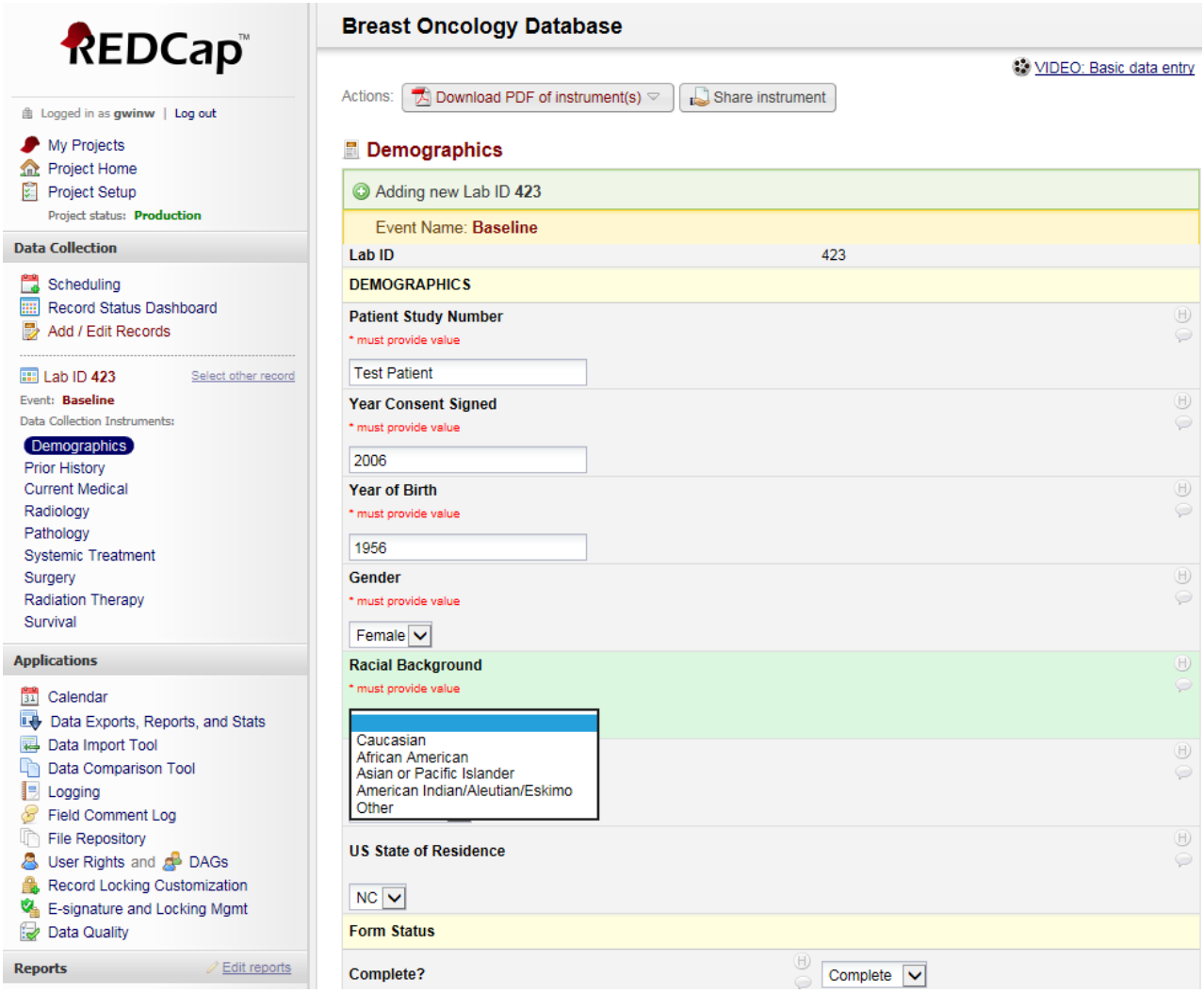
Pathologic and Clinical Annotation Database

A clinical annotation database titled the Breast Oncology Database has been established to complement the procured SPORE sample characteristics and annotated pathology data. This Breast Oncology Database is an offsite clinical annotation database created in collaboration with Cedar-Sinai Medical Center. This database is built on the REDCap platform and includes the full functionality of REDCap. This database is housed on a server at Cedars-Sinai Medical Center (CSMC) and is accessible on the web via https (no VPN required). This database adheres to CSMC Enterprise Information Services (EIS) research database security standards. The Breast Oncology Database consists of: 9 Baseline forms and 4 Follow-up forms with a total of 779 possible entry fields. The clinical data elements collected include: demographics, radiographic findings (mammography, U/S, MRI, CT, PET), pathologic diagnosis (histology, IHC, and FISH), surgical and radiation treatment received, systemic treatment received (dates and dosing regimens), and survival data.

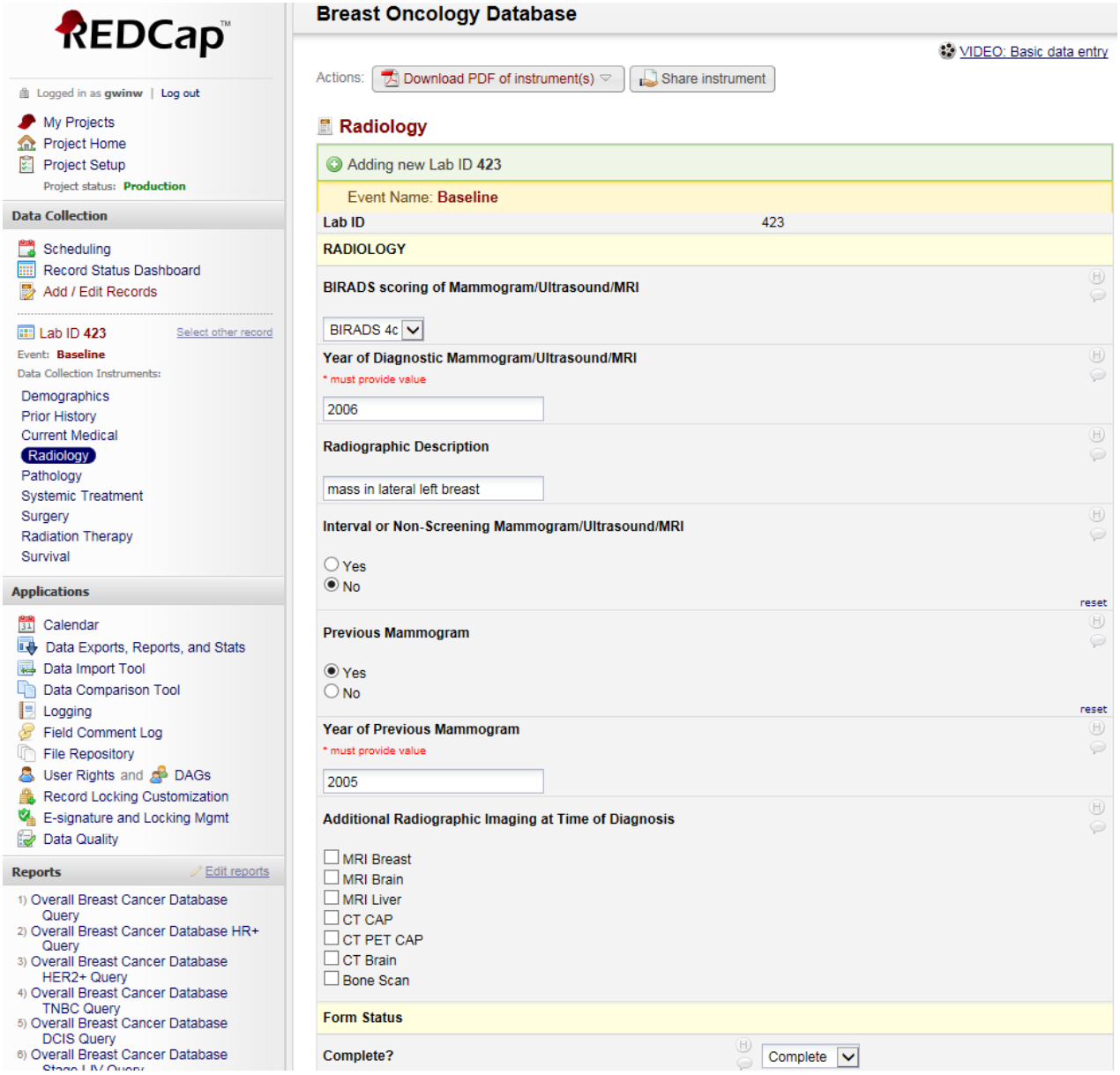
To date we have annotated 197 cases of DCIS and DCIS/IDC contained in the SPORE tissue bank, this includes the 78 cases sent to Dr. Hannon's lab for genomic evaluation. Below is a clinical summary of the clinical cases related to the 78 tumor samples sent to Dr. Hannon's lab:

- Median Age at enrollment: 55.4 yrs (range: 35-89)
- Menopausal status: Pre-menopause: 37.2%, Post-meno: 62.8%
- Prior IDC or DCIS: 3.9% prior IDC, 2.6% prior DCIS
- HR and HER2 Receptor status:
  - ER/PR+, HER2+ : 6 cases
  - ER/PR+, HER2- : 54 cases
  - ER/PR-, HER2+ : 6 cases
  - ER/PR-, HER2- : 9 cases
- Received Neoadjuvant or adjuvant systemic therapy: 50 cases
- Recurrences of Breast Ca since SPORE enrollment: 15 cases
- Death since SPORE enrollment: 6 cases

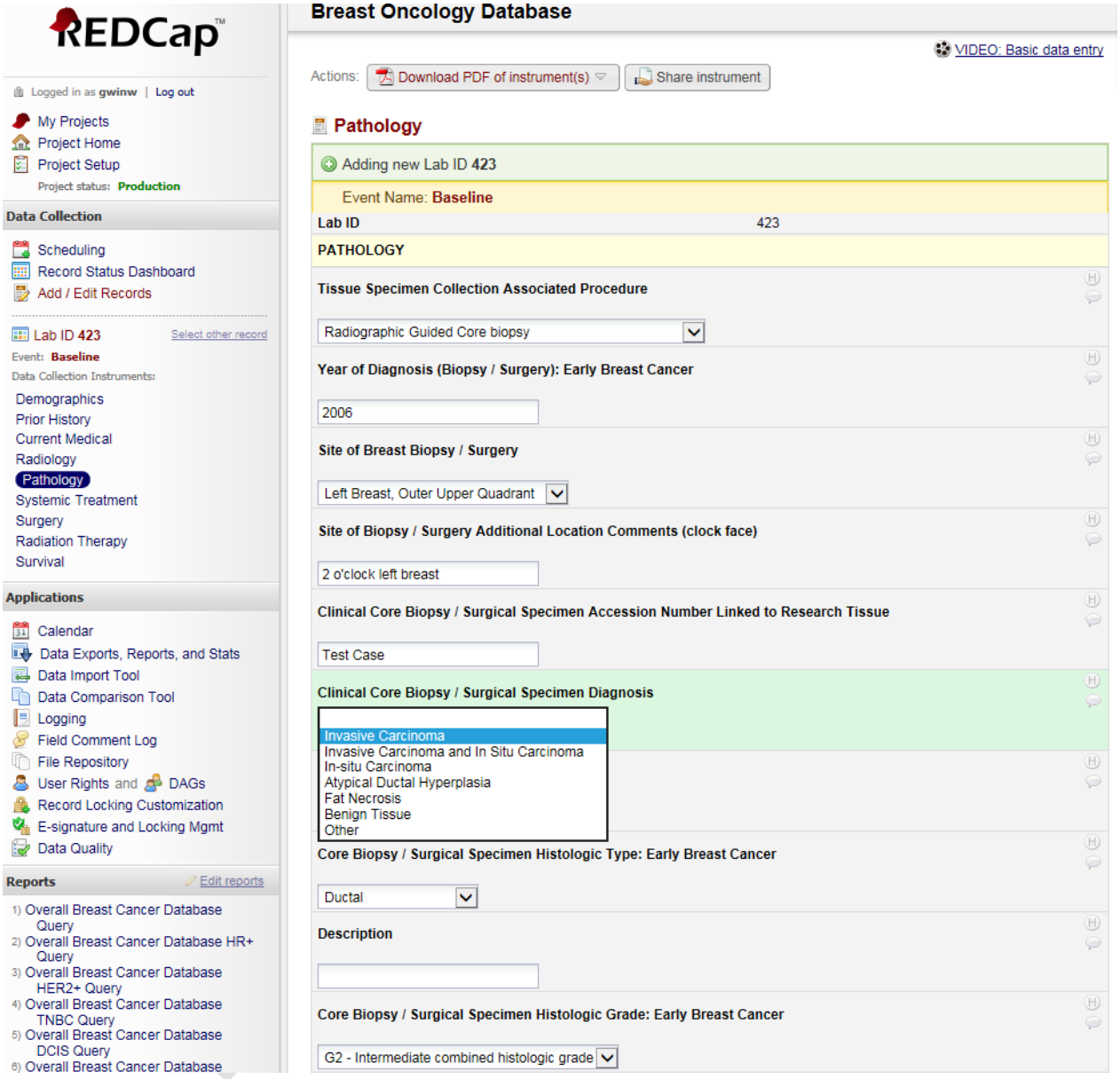
Demographics



Radiology


REDCap
Logged in as gwinw | Log out
My Projects
Project Home
Project Setup
Project status: Production
Data Collection
Scheduling
Record Status Dashboard
Add / Edit Records
Lab ID 423
Select other record
Event: Baseline
Data Collection Instruments:
Demographics
Prior History
Current Medical
Radiology
Pathology
Systemic Treatment
Surgery
Radiation Therapy
Survival
Applications
Calendar
Data Exports, Reports, and Stats
Data Import Tool
Data Comparison Tool
Logging
Field Comment Log
File Repository
User Rights and DAGs
Record Locking Customization
E-signature and Locking Mgmt
Data Quality
Reports
Edit reports
1) Overall Breast Cancer Database Query
2) Overall Breast Cancer Database HR+ Query
3) Overall Breast Cancer Database HER2+ Query
4) Overall Breast Cancer Database TNBC Query
5) Overall Breast Cancer Database DCIS Query
6) Overall Breast Cancer Database
Breast Oncology Database
VIDEO: Basic data entry
Actions:
Download PDF of instrument(s)
Share instrument
Radiology
Adding new Lab ID 423
Event Name: Baseline
Lab ID 423
RADIOLOGY
BIRADS scoring of Mammogram/Ultrasound/MRI
BIRADS 4c
Year of Diagnostic Mammogram/Ultrasound/MRI
* must provide value
2006
Radiographic Description
mass in lateral left breast
Interval or Non-Screening Mammogram/Ultrasound/MRI
Yes
No
reset
Previous Mammogram
Yes
No
reset
Year of Previous Mammogram
* must provide value
2005
Additional Radiographic Imaging at Time of Diagnosis
MRI Breast
MRI Brain
MRI Liver
CT CAP
CT PET CAP
CT Brain
Bone Scan
Form Status
Complete?
Complete

Pathology



- Include: ER/PR/HER2, T N M- Stage, pathologic and clinical staging options
- Research Pathology description and concurrence
- Related genomic mutations (BRCA1/2, etc.) detected in clinical work-up

Searching of the Database: Predefined Queries

8	SPORE Breast Cancer Database HR+ Query	View Report	Export Data	Stats & Charts	Edit	Copy	Delete
9	SPORE Breast Cancer Database HER2+ Query	View Report	Export Data	Stats & Charts	Edit	Copy	Delete
10	SPORE Breast Cancer Database TNBC Query	View Report	Export Data	Stats & Charts	Edit	Copy	Delete
11	SPORE Breast Cancer Database DCIS Query	View Report	Export Data	Stats & Charts	Edit	Copy	Delete
12	SPORE Breast Cancer Database Stage I-IV Query	View Report	Export Data	Stats & Charts	Edit	Copy	Delete

Name of Report: SPORE Breast Cancer Database All Cases Query

STEP 1

User Access: Choose who can view this report

(•) **All users** – OR – ( ) **Custom user access** (Choose specific users, roles, or data access groups who will have access)

STEP 2

Fields to include in report Add all fields from selected instrument: -- choose instrument --

Field 1	lab_id "Lab ID"	Instrument: **Demographics**
Field 2	subject_number "Patient Study Number"	Instrument: **Demographics**
Field 3	year_of_birth "Year of Birth"	Instrument: **Demographics**
Field 4	racial_background "Racial Background"	Instrument: **Demographics**
Field 5	ethnicity "Ethnicity"	Instrument: **Demographics**
Field 6	prev_carc_situ "Previous Carcinoma/In-situ	Instrument: **Prior History**
Field 7	menstrual_status "Menstrual Status (at time	Instrument: **Prior History**
Field 8	clinical_core_biopsy_diag "Clinical Core Bio	Instrument: **Pathology**
Field 9	year_diagnosis_biopsy "Year of Diagnosis (	Instrument: **Pathology**
Field 10	hormone_receptor_1 "Hormone Receptor S	Instrument: **Pathology**
Field 11	hormone_receptor_3 "Hormone Receptor S	Instrument: **Pathology**
Field 12	her2_status_1 "HER2 Status: Early Breast (	Instrument: **Pathology**
Field 13	acrb_acsd "Additional Clinical Specimen Di	Instrument: **Pathology**
Field 14	acrb_acserprs "Additional Clinical Specimer	Instrument: **Pathology**
Field 15	acrb_acsher2s "Additional Clinical Specime	Instrument: **Pathology**
Field 16	clincial_staging "Pathologic Staging"	Instrument: **Pathology**
Field 17	c_stage_diag "Clinical Staging at Time of Di	Instrument: **Pathology**
Field 18	neo_sys_therapy "Neoadjuvant Systemic Th	Instrument: **Systemic Treatment**
Field 19	adjuv_sys_therapy "Adjuvant Systemic Ther	Instrument: **Systemic Treatment**
Field 20	sys_palli_therapy "Systemic / Palliative The	Instrument: **Systemic Treatment**

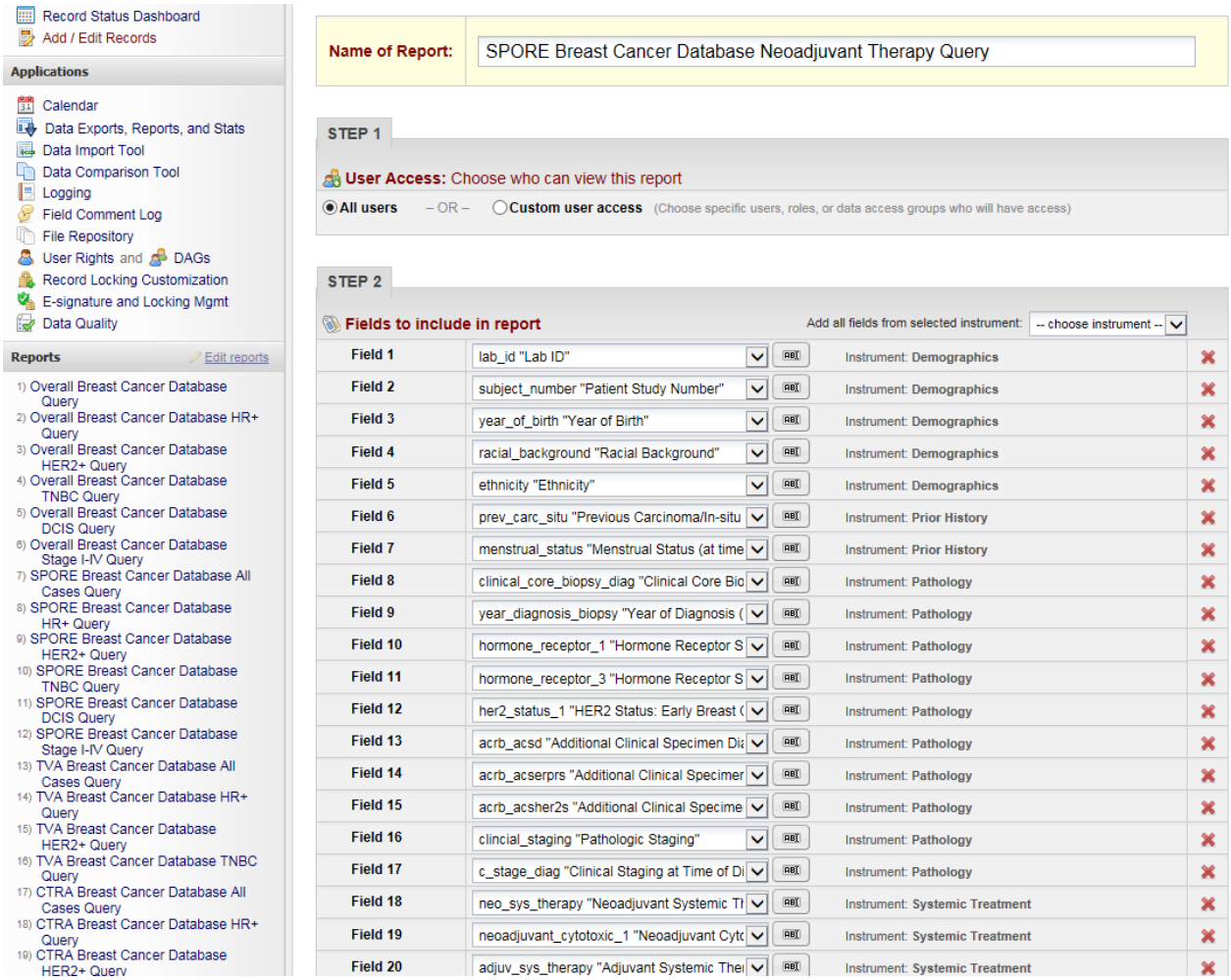
Searching of the Database: Predefined Queries Results

Overall Breast Cancer Database DCIS Query

Enable floating table hea

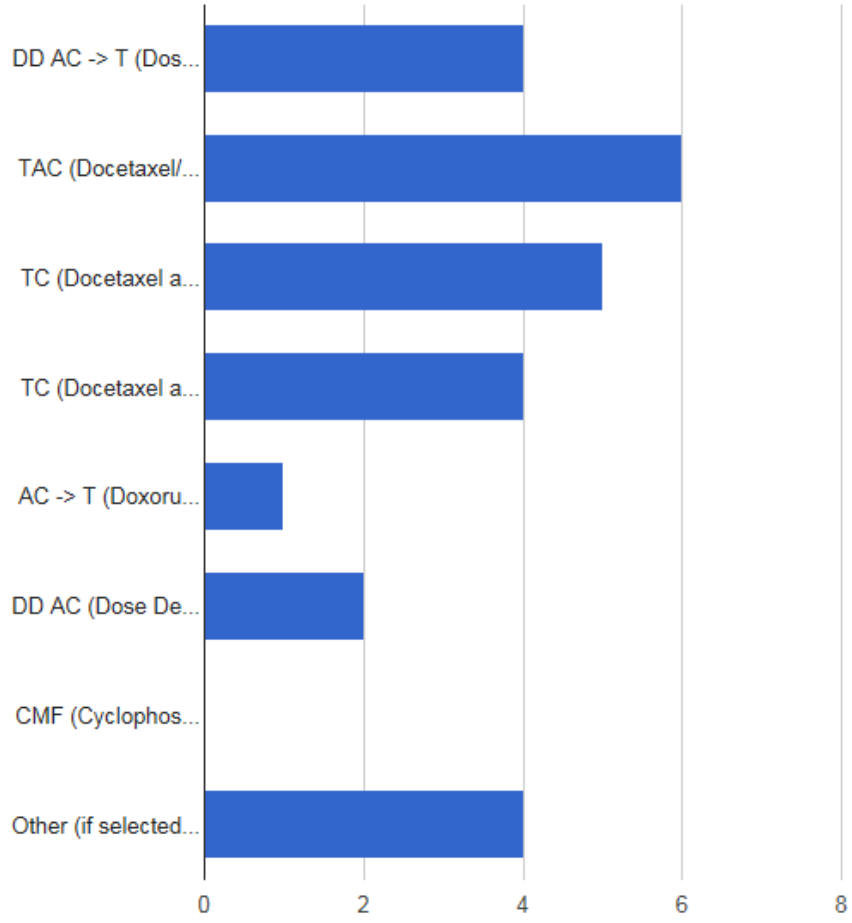
Lab ID (lab_id)	Event Name (redcap_event_name)	Year of Birth (year_of_birth)	Racial Background (racial_background)	Ethnicity (ethnicity)	Previous Carcinoma/In-situ Carcinoma History (prev_carc_situ)	Menstrual Status (at time of diagnosis) (menstrual_status)	Clinical Core Biopsy / Surgical Specimen Diagnosis (clinical_core_biopsy_diag)	Year of Diagnosis (Biopsy / Surgery): Early Breast Cancer (year_diagnosis_biopsy)	Hormone Receptor Status (ER/PR): Early Breast Cancer (hormone_receptor_1)	Hormone Receptor Status (ER/PR): Early Breast Cancer (hormone_receptor_3)	HER2 Status: Early Breast Cancer (her2_status_1)	Additional Clinical Specimen Diagnosis (acrb_acsd)	Additional Clinical Specimen ER/PR Status (acrb_acserprs)	Non-Breast BioBiopsy / Surgical Specimen HER2 status (non_breast_bb_ss_her2)	Pathologic Staging (clincial_staging)
1	Baseline	1958	Other (6)		No prior history of breast carcinoma or in-situ breast carcinoma (4)	Postmenopausal (2)	In-situ Carcinoma (2)	2005	ER(-) (2)	PR(-) (4)	HER2 non-amplified / not over-expressed (1)				Stage 0 (1)
2	Baseline	1971	Caucasian (1)	Non-Hispanic (2)	Other (3)	Premenopausal (1)	In-situ Carcinoma (2)	2005							Stage 0 (1)
3	Baseline	1933	African American (2)	Non-Hispanic (2)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Postmenopausal (2)	In-situ Carcinoma (2)	2006	ER(+) (1)	PR(+) (3)					Stage 0 (1)
4	Baseline	1967	Caucasian (1)	Non-Hispanic (2)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Premenopausal (1)	In-situ Carcinoma (2)	2007	ER(+) (1)	PR(-) (4)					Stage 0 (1)
5	Baseline	1940	Caucasian (1)	Non-Hispanic (2)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Postmenopausal (2)	In-situ Carcinoma (2)	2007	ER(+) (1)	PR(-) (4)					Stage 0 (1)
16	Baseline	1966	Caucasian (1)	Hispanic (1)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Premenopausal (1)	In-situ Carcinoma (2)	2005	ER(-) (2)	PR(-) (4)					Stage 0 (1)
17	Baseline	1949	Caucasian (1)	Non-Hispanic (2)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Postmenopausal (2)	In-situ Carcinoma (2)	2006	ER(+) (1)	PR(+) (3)					Stage 0 (1)
18	Baseline	1962	African American (2)	Non-Hispanic (2)	No prior history of breast carcinoma or in-situ breast carcinoma (4)	Premenopausal (1)	In-situ Carcinoma (2)	2007	ER(+) (1)	PR(+) (3)					Stage 0 (1)
20	Baseline	1958	Caucasian (1)	Non-Hispanic (2)		Premenopausal (1)	In-situ Carcinoma (2)	2004	ER(+) (1)	PR(+) (3)		In Situ Carcinoma (3)	Not evaluated (4)		Stage 0 (1)
47	Baseline	1933	African American (2)	Non-Hispanic (2)	In-situ breast carcinoma (2)	Postmenopausal (2)	In-situ Carcinoma (2)	2014	ER(-) (2)	PR(-) (4)					Stage 0 (1)

Searching of the Database: New Database Queries – SPORE cases with neoadjuvant therapy



Searching of the Database: SPORE cases with neoadjuvant therapy Query – Results

Counts/frequency: DD AC -> T (Dose Dense Doxorubicin/Cyclophosphamide followed by Paclitaxel) (4, 15.4%), TAC (Docetaxel/Doxorubicin/Cyclophosphamide) (6, 23.1%), TC (Docetaxel and Cyclophosphamide) (5, 19.2%), TC (Docetaxel and Carboplatin) (4, 15.4%), AC -> T (Doxorubicin/Cyclophosphamide followed by Paclitaxel) (1, 3.8%), DD AC (Dose Dense Doxorubicin/Cyclophosphamide) (2, 7.7%), CMF (Cyclophosphamide/Methotrexate/Florouracil) (0, 0.0%), Other (if selected provide free text below) (4, 15.4%)

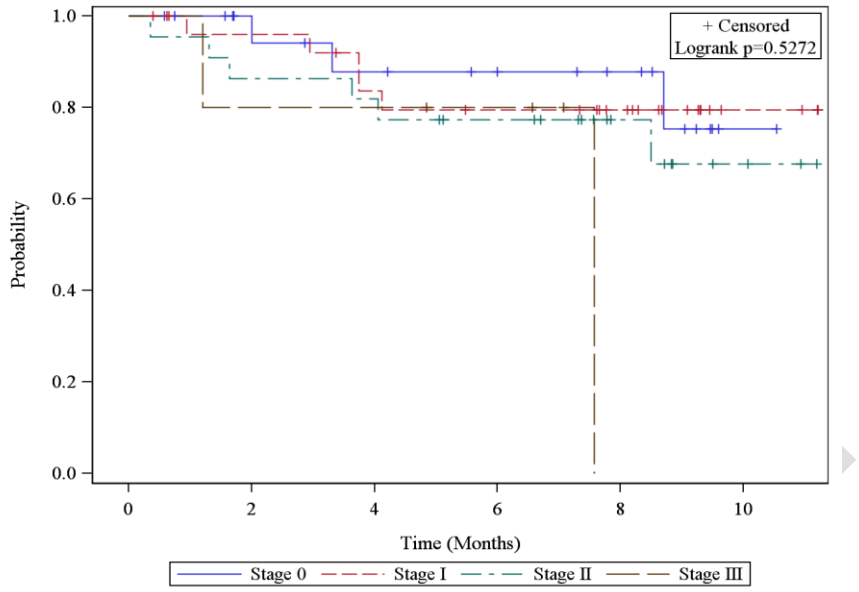


Database Statistical Analysis Options

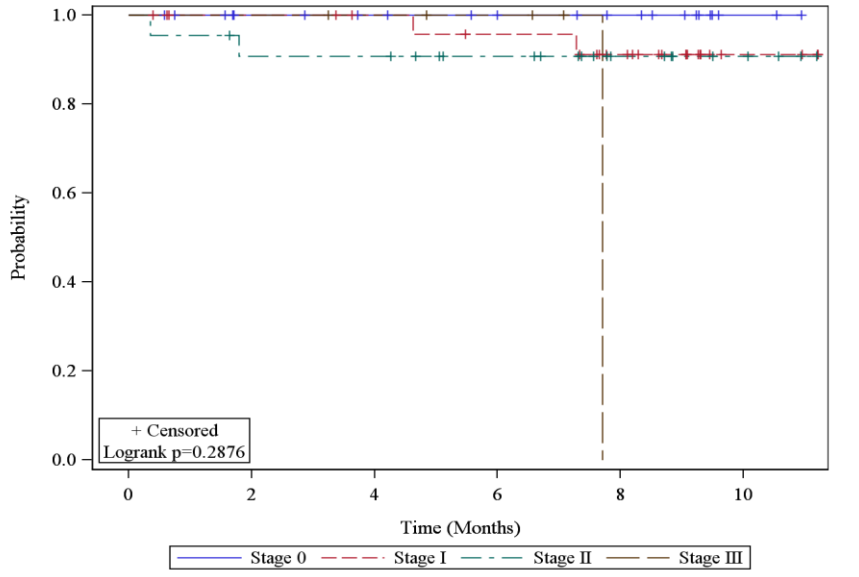
	Download Syntax & Data
Microsoft Excel You may download the survey results in CSV (comma-separated) format, which can be opened in Excel. You have the choice of downloading the data either with the full headers and answer labels or just with the answer codes (i.e. raw data). *NOTE: If you are using a version of Microsoft Excel prior to Excel 2007, due to limitations the data will only be read to 255 columns when opened.*	EXCEL CSV Labels · EXCEL CSV Raw Send file?
SPSS Statistical Analysis Software Instructions: Download and save all 3 files on the right to a common location. First, double-click on the Pathway Mapper (.bat) file, which will run quickly and invisibly. (If you are not using a Windows operating system, such as Mac or Linux, please see the *Additional Instructions*.) Now double-click on the *.sps file, which will open SPSS. When the file is loaded and displayed, choose Run-->All from the top menu options. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc. Additional instructions	SPSS · DATA CSV Pathway Mapper Send file?
SAS Statistical Software Instructions: Download and save all 3 files on the right to a common location. First, double-click on the Pathway Mapper (.bat) file, which will run quickly and invisibly. (If you are not using a Windows operating system, such as Mac or Linux, please see the *Additional Instructions*.) Now double-click on the *.sas file, which will open SAS. When the file is loaded and displayed, choose Run (or Run-->Submit) from the top menu options. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc. Additional instructions	SAS · DATA CSV Pathway Mapper Send file?
R Statistical Software Instructions: Use command read.csv('filename') to read in data file.	R · DATA CSV Send file?
STATA Analysis and Statistical Software Instructions: Download both files to common location and double-click on *.do file. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc.	STATA · DATA CSV Send file?

Database Statistical Analysis – Disease Free Survival and Overall Survival Curves of Annotated Cases

Kaplan-Meier Disease Free Survival curves from Breast Oncology Database annotation of tissues sent to Dr. Hannon's Lab:



Kaplan-Meier Overall Survival curves from Breast Oncology Database annotation of tissues sent to Dr. Hannon's Lab:



Duke Pathology Archives

Between 600 and 700 breast surgical procedures are performed at Duke University Medical Center each year, and even though the number of surgeries has been relatively steady, a much smaller percentage of cancers have been cryopreserved and banked in recent years. Fortunately, we currently have over 50,000 samples have been archived following formalin fixation and paraffin embedding.

Utilizing Duke's Breast Data Mart

We have identified the Duke Breast Data Mart as a valuable resource. The Breast Research Data Mart is a component of the Research Management Data Mart (RMDM) which allows:

- Data integration from MaestroCare (EPIC), IRB (eIRB), CTMS (eResearch), Finance (SAP and PDC), HR (Faculty and Staff) and other base operational systems to present a complete view of research activities
- Reports and Dashboards (such as Enrollment, Protocol, Finance etc.) to be created to assist with monitoring of individual projects.

A semantic layer of the Business Objects Universe has been created to provide end users the ability to intuitively query and retrieve data consistently. Schematic below shows the current state and long term vision for RMDM.

Utilizing the Breast Data Mart, we have identified nearly 1,000 additional patients with clinical and pathologically confirmed diagnosis of ductal carcinoma in situ (DCIS).

We have drafted a retrospective data analysis protocol entitled "The Natural History of Ductal Carcinoma In Situ and the Identification of Clinical Outliers" and plan to submit this to the Duke IRB in September 2016. This protocol would allow us access to clinically annotate the medical records of the previously identified nearly 1,000 DCIS patients seen at Duke from 2004 to 2013. Clinical annotation of DCIS patients may allow us to observe trends in molecular subtypes of DCIS (i.e., hormone sensitive, HER2 over-expression, etc.), treatment modalities, progression of disease, and survival outcomes. In addition, we plan to file a waiver of consent with the Duke IRB in order to identify patients treated at Duke with DCIS. Once patients with a history of DCIS are identified, we may approach them to request consent to use excess tissue, previously collected for clinical pathology, for research purposes.

Dr. Greg Hannon of Cancer Research UK Cambridge Institute will provide his research accomplishments creating libraries and sequencing these samples via a separate report.

Opportunities for training and professional development

Nothing to report (not intended for training).

Results disseminated to communities of interest

Nothing to report.

Plan for next reporting period

We plan to continue to clinically annotate the Duke frozen tissue breast bank. We will then extend our clinical annotation to the archival breast samples from Duke pathology, consisting of formaldehyde fixed paraffin embedded tissue after identifying these from the Duke pathology reports.

These samples will be identified and initial attempts to assemble them into a validation cohort for subsequent markers.

4. Reportable outcomes

Dr. Joe Geradts continues to perform the tissue microdissection and accomplished the following during this past year:

1. Confirmed the presence of DCIS and related pathologic lesions in frozen sections prepared in the Hannon lab from biopsies provided by the Duke group.

2. Annotated a large number of images from the breast core biopsies sent to the Hannon lab for microdissection. The exact number of cases and the range of annotated regions (up to ~60 per case) should be included in the Hannon report. The different regions of interest were numerically coded as follows:

   1 = invasive carcinoma
   2 = DCIS
   3 = benign epithelium
   4 = normal epithelium
   5 = stroma adjacent to invasive carcinoma
   6 = stroma adjacent to DCIS
   7 = stroma away from DCIS/invasive ca
   8 = inflammatory focus
   9 = other

   For each type of annotation region, a/b/c/d indicate individual foci within that region.

3. Confirmed that a large number of annotated regions were correctly microdissected. The exact number may be included in the Hannon report.

4. Several face-to-face meetings, telephone conversations, and extensive e-mail correspondence with the Hannon and Duke groups, given separate physical locations.

Changes in use or care of human subjects

Nothing to report.

Products

Nothing to report.

5. Participants & Other Collaborating Organizations

Individuals worked on the project

Name: **H. Kim Lyerly**
Project Role: Partnering PI – provides strategic insight into DCIS biology and clinical outcomes, and continues to engage thought leaders to support ongoing activities interrogating DCIS biology. Coordinates efforts to identify appropriate tumor samples for analysis from the Duke SPORE tissue bank, the clinical annotation of the samples in both the SPORE tissue bank, and the general Duke breast tissue bank, and administration of the the general tissue microdissection activities that will be directly supervised by Dr. Geradts and his team.
Nearest person month worked: 2.05 CM

Name: **Joseph Geradts**
Project Role: Co-investigator – microdissection of tissues and interpretation of stained tissue sections
Nearest person month worked: 1.2 CM

Name: **Katherine Kalinowski**
Project Role: Associate in Research – clinical data abstraction and annotation into secure DCIS clinical annotation database
Nearest person month worked: 4.8 CM

Name: **Amy Hobeika**
Project Role: Regulatory Administrator – ensures all necessary regulatory documentation is submitted and maintained
Nearest person month worked: 1.32 CM

Name: **Qing Cheng**
Project Role: Co-investigator – provide support for RNAseq analysis and bioinformatics coordination of genomics data with the clinical outcomes data with the clinical database
Nearest person month worked: 3.18 CM

Name: **Kimberly Egler**
Project Role: Research Project Manager – oversees administrative and financial activities
Nearest person month worked: 1.92 CM

Name: **Delila Serra**
Project Role: Laboratory Research Analyst – performs immunohistochemical assays on breast tumor tissue
Nearest person month worked: 6 CM

Name: **Karrie Comatas**
Project Role: Laboratory Research Analyst – performs immunohistochemical assays on breast tumor tissue
Nearest person month worked: 6 CM

Change in active support since last report

Dr. Lyerly's effort changed from 2.52 to 2.05 CM; as the grant moved into year 2, research staff were trained on their tasks and less time was needed to oversee efforts and provide guidance. Similarly, Kimberly Egler's effort went from 2.4 to 1.92 CM.

Other organizations involved as partners

Cancer Research UK Cambridge Institute – collaboration to perform the RNA sequencing and the DNA sequencing and have laser microdissected 35 patients, as detailed in the grant application.

Cedars Sinai Medical Center (CSMC) – storage and management of secure DCIS database; as detailed in the grant application.

6. Conclusions

Annotation will continue and identification and annotation of the validation cohort from the Duke pathology archives will continue.

7. Appendices

Tissue Use Protocol: Molecular Framework of Early Breast Cancer
Duke IRB Continuing Review Approval

Tissue Use Protocol

Molecular Framework of Early Breast Cancer

Study Investigators

Principal Investigator: H. Kim Lyerly, MD
Duke University Medical Center
Department of Surgery
MSRB1 Research Dr Box 2606
Durham, NC 27710

Sub-Investigators: Michael A. Morse, MD
Rajesh Dash, MD
Shannon McCall, MD
William Gwin III, MD
Duke University Medical Center
Greg Hannon, PhD
Cold Spring Harbor

Statistician: Steven Piantadosi, MD, PhD
Cedars Sinai

Regulatory Coordinator: Amy Hobeika, PhD
Duke University Medical Center

Study Location: Duke University Medical Center
FWA FWA00009025

Research Summary

1. Protocol Title: Molecular Framework of Early Breast Cancer

2. Purpose of the Study: The objective of the study is to perform molecular profiling of early breast cancer using human breast cancer specimens that have been banked and will be banked in the future in the Breast SPORE tissue bank (Protocol # Pro00014678), the DUHS Biospecimen Repository and Processing Core (BRPC) Facility (Protocol #Pro00035974), the DOD TVA tissue bank (Protocol #Pro00045965), and the DOD CTRA tissue bank (Protocol #Pro00044981).

3. Background & Significance: Breast cancer remains a leading cause of death amongst women worldwide. There has been a dramatic rise in the incidence of ductal carcinoma in situ (DCIS) breast cancer in the US which is at least partially due to the use of screening mammography [1]. Currently, therapy for DCIS involves surgical resection with or without radiation and sometimes adjuvant hormone therapy [2-4]. Although Page et al. [5] have shown that overall there is a nine-fold increased likelihood for women with DCIS to develop invasive breast carcinoma, there are few biomarkers to predict behavior in individual patients with DCIS [6,7].

Early discovery of a DCIS lesion should be a lifesaver, it should not be the beginning of needless treatment and suffering for half the people diagnosed. Our objective is to undertake extensive molecular profiling of DCIS and early invasive disease, identifying signatures and patterns that can be used to predict long term, and short term, outcomes and help guide choices among treatment options for early breast cancer.

<u>Exome Capture</u>

The common perception of DCIS is that it is the precursor to invasive disease. This however, has never been shown, nor is there any understanding of how DCIS lesions within the same patient and across patients are related to each other. We will use exome capture to establish copy number variations and somatic variations of DCIS lesions to determine if each lesion is a result of an independent event or a founder effect, and how evolutionary connected different pathological grades of DCIS are. We will also identify shared variations between invasive cells and DCIS, to finally establish if invasive disease is a result of DCIS progression. Shared variations between patients will provide possible biomarkers for predictive power over patient outcome.

<u>Transcriptional analysis of early breast cancer</u>

Using the latest in laser capture and high-throughput technology, we will isolate DCIS lesions and neighboring cells from biopsies. Making use of RNAseq methods specific for low input samples we will provide a transcriptional multilayered landscape. We will look for clustered, and unique, expression changes across hundreds of patient biopsies, covering a range of DCIS and early breast cancer subtypes, and where possible associated normal tissue. This will allow us to generate hypotheses for molecular markers.

<u>The role of stroma in DCIS outcomes</u>

Current literature suggests that the stromal compartment surrounding invasive breast cancers, and even DCIS, could play a role in the progression of the disease. The break from DCIS to invasive carcinoma could be a result not of the cancer cells them selves changing, but of the stromal compartment shifting to accommodate the expansion of tumor cells. It is not yet know what stromal cell types, if any, play a role in the development of early invasive breast cancer. We will carry out histology of a subset of samples. Using cell type specific markers, we will visualize the cellular community surrounding both the DCIS and the early invasive carcinoma.

4. Design & Procedures:
We plan to evaluate the biology of human DCIS through genomic analysis including tissue laser capture, transcriptome sequencing (RNA seq), exome sequencing, and immunohistochemistry.

We plan to obtain the breast and associated normal tissues either from biopsies banked by the DUHS Biospecimen Repository, Processing Core (BRPC) Facility or the Breast SPORE Tumor Bank, the DOD TVA tissue bank, and/or the DOD CTRA tissue bank. We are interested in studying paraffin embedded tissue, fresh tissue, and/or frozen early breast cancer specimens and benign breast tissue.

The Duke SPORE Breast tumor bank contains 7,000 core needle samples from 1,700 patients. These samples are pathologically confirmed and annotated.

The DUHS Biospecimen Repository and Processing Core (BRPC) Facility is a centralized biospecimen collection and processing infrastructure with standardized protocols with a robust quality assurance program to assure sample integrity, data security and patient safety. The facility is responsible for maintenance of the Breast SPORE tumor bank collection and also prospectively collects breast tumor specimens for the Breast site-based group at DUMC.

The DOD TVA tissue bank) archives core biopsies and blood from subjects with mammographically detected breast abnormalities that are undergoing breast biopsy. These samples are pathologically confirmed and annotated.

The DOD CTRA tissue bank archives tumor tissue and blood from subjects with a history of breast cancer that are undergoing medically indicated surgical or biopsy procedures related to their breast cancer. These samples are pathologically confirmed and annotated.

The tissue provided for this use protocol will contain study ID numbers based on the tissue bank from which they are received. Study investigators from Duke will have access to review clinical data related to tissues used in this protocol, including access to patient PHI, including name and MRN.

Brief description of assays to be prioritized. ‘

Genomic Analysis (Exome sequencing)

We will perform tissue laser capture on breast tumor samples followed by in depth profiling methods to define the genetic alterations associated with early breast cancer

and answer critical questions on how lesions from the same person are associated with each other, and how clonal evolution contributes to the progression to invasive cancer.

Exome sequencing provides an effective means of obtaining both copy number and somatic mutation information for studies with a large number of patients (and many lesions per patient) compared to whole genome sequencing. We plan to use the ExomeCNV algorithm [8] to call copy number aberrations in breast samples as this algorithm was designed for tumor-normal comparison data.
We plan to identify variants using the GATK toolkit, which shows the highest precision and recall for identifying single nucleotides variants (SNVs) and small indels. Copy number variation (CNV) and SNV data from multiple DCIS and IDC lesions will be used to determine the relatedness between samples.

Transcriptome Sequencing (RNA Sequencing)

We will isolate DCIS lesions and neighboring cells from the same patient biopsies using high-throughput technology methods including the laser capture. Making use of RNAseq methods specific for low input samples we will look for clustered, and unique, expression changes in patient biopsies, covering a range of DCIS and early breast cancer subtypes, and associated normal tissue to identify transcriptional changes that lead to DCIS and early invasive cancers. This data will be used to generate hypotheses for molecular markers.

We will also use RNAseq to determine the role of stroma in early breast cancer outcomes. For this analysis, high quality RNA seq libraries will be made using small sample inputs captured by laser microscopy. The RNAseq libraries will be mapped to the human genome and RNA-seQC used to evaluate the transcript coverage rates, duplication rates, GC content bias, and exonic/intronic alignment fractions. Expression profiles will be analyzed to see if the pathological grades of DCIS cluster together and if samples from the same patient cluster together.

Validating biomarkers of DCIS progression (Immunohistochemistry)

Information gathered from genomic analysis will be used to generate hypotheses regarding potential prognosis makers. Assays will be optimized and applied to paraffin sections from DCIS lesions. Immunostains will be manually scored by a pathologist. Evaluation of the IHC stains will take into account subcellular localization, percentage of staining cells, staining pattern (diffuse, mosaic, patchy, focal) and staining intensity.

5. Selection of Subjects: We will identify up to 400 individual specimens (multiple may be from the same donor) in total to complete the proposed studies from the following biospecimen repositories: DUHS Biospecimen Repository, Processing Core (BRPC) biobank or the Breast SPORE Tumor Bank, the DOD TVA tissue bank, and/or the DOD CTRA tissue bank. We will request treatment refractory breast cancer specimens of each subtype (ER+, HER2+, triple negative).

6. Subject Recruitment & Compensation: We are obtaining archived tumor tissue samples and tissues which may be collected in the future from the Breast SPORE Tissue Bank and from the DUHS Biospecimen Repository and Processing Core (BRPC) Facility (Director Shannon McCall). Samples collected under the DOD TVA and DOD CTRA tissue banks may also be used. There will not be compensation to the subjects.

7. Consent Process: Subjects have been consented or will be consented for tissue banking and study under the Duke Breast SPORE grant (Pro00014678),the DUHS Biospecimen Repository and Processing Core (BRPC) Facility protocol, the DOD TVA tissue bank (Protocol #Pro00045965), or the DOD CTRA tissue bank (Protocol #Pro00044981). They will not be re-consented in this study.

8. Subject's Capacity to Give Legally Effective Consent: Subjects will not be re-consented in this study.

9. Study Interventions: There is no study intervention.

10. Risk/Benefit Assessment: Because the tissue has already been obtained, we do not expect any risks to the patients from whom the tissue specimens were obtained other than a small risk that patient confidentiality could be compromised by review of clinical data. There are no immediate benefits to patients. Results of these studies may benefit future patients with breast cancer.

11. Costs to the Subject: There is no cost to the subject.

12. Data Analysis & Statistical Considerations:

Sample Size and Power: The power of RNA sequencing for detecting transcriptional changes that leads to DCIC and earlier invasive cancers is estimated with RNASeqPower package in R [9]. With the depth of sequencing of 20, the coefficient of variation of 0.75, the significance level of 0.05, and the effect size of 1.5, the minimal sample sizes of 59 and 79 for each group are required to achieve the statistical power of 80% and 90% respectively. The power to detect CNVs is estimated with the Online Exome Power Calculator (exomepower.ssg.uab.edu). With the number of mutations m = 300, total number of genes M = 20,000, sensitivity of detecting mutations Ps = 0.8, the mutation probability equals the genome-wide average w = 1, and genetic heterogeneities of 0.05, the minimal sample size of 211 is required to achieve the 80% power for detecting mutations. There will be enough specimens for each group, since we have up to 400 individual specimens in this study.

Data Analysis: While ExomeCNV software will be used for CNV detection, we plan to use our own transcriptome analysis pipeline for RNA-seq, which uses the Tophat (http://tophat.cbcb.umd.edu/) software package for performing gapped alignments against the reference genome, DESeq for detecting differential gene expression, and Cuffsuite[10] for detecting differential isoform expression. These tools provide a list of differentially expressed genes and isoforms and provide p-values, FDR, q-value, and normalized expressions values (normalized read counts in the case of Cuffdiff). We typically use a combination of FDR (<0.05), log fold change, and a noise filter with a minimum number of reads or FPKM for a gene or isoform in all the samples to identify differentially expressed genes and isoforms. After we have both the CNV, somatic mutation, and the differentiated expressed genes and isoforms, we will construct eQTL mapping with linear regression to dissect the genetic basis of gene expression.
PPI networks in BioGRID and other online functional databases (e.g. Gene Ontology (GO), KEGG, and Enzyme Commission (EC) will be used to explore the biological mechanisms of the identified genes. Prediction models such as support vector machines and logistic regression will be used to identify the potential biomarkers and evaluate their

predict powers.

13. Data & Safety Monitoring: There will be no data and safety monitoring of the study. The protocol will be conducted in accordance with the protocol submitted to and approved by the USAMRMC ORP HRPO and will not be initiated until written notification of approval of the research project is issued by the USAMRMC ORP HRPO.

14. Privacy, Data Storage & Confidentiality: All records pertaining to the identity of samples obtained from the tissue repository will be maintained as private and confidential. Patient information accessed for this study will be kept in a password-protected database on a server managed by the Duke Department of Surgery IT support team that will be accessible only to Dr. Lyerly and the study key personnel within Duke. All records pertaining to the identity of participants in the tissue repository will be maintained as private and confidential. PHI will not be provided to personnel outside of Duke. RNA sequencing data will be generated at Cold Spring Harbor Laboratories by Dr. Greg Hannon. Deidentified tissue will be sent to Dr. Hannon at Cold Spring Harbor for RNA sequencing. Dr. Hannon will not receive any PHI and will not seek to obtain PHI. Once RNA sequencing is completed by Dr. Hannon, data will be sent to Dr. Lyerly.

Key Personnel
PI: H. Kim Lyerly, MD
Sub-Investigators: Qing Cheng, PhD, Rajesh Dash, MD, Greg Hannon, PhD, William Gwin III, MD

References
1. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA.* 1996;275(12):913–918.
2. Czerniecki BJ, Koski GK, Koldovsky U, Xu S, Cohen PA, Mick R, Nisenbaum H, Pasha T, Xu M, Fox KR, Weinstein S, Orel SG, Vonderheide R, Coukos G, DeMichele A, Araujo L, Spitz FR, Rosen M, Levine BL, June C, Zhang PJ. Targeting HER-2/neu in early breast cancer development using dendritic cells with staged interleukin-12 burst secretion. *Cancer Res.* 2007;67(4):1842–1852.
3. Gonzalez RJ, Buzdar AU, Fraser Symmans W, Yen TW, Broglio KR, Lucci A, Esteva FJ, Yin G, Kuerer HM. Novel clinical trial designs for treatment of ductal carcinoma in situ of the breast with trastuzumab (herceptin) *Breast J.* 2007;13(1):72–75.
4. Yen TW, Kuerer HM, Ottesen RA, Rouse L, Niland JC, Edge SB, Theriault RL, Weeks JC. Impact of randomized clinical trial results in the national comprehensive cancer network on the use of tamoxifen after breast surgery for ductal carcinoma in situ. *J Clin Oncol.* 2007;25(22):3251–3258.
5. Page DL, Dupont WD, Rogers LW, Jensen RA, Schuyler PA. Continued local recurrence of carcinoma 15–25 years after a diagnosis of low grade ductal carcinoma in situ of the breast treated only by biopsy. *Cancer.* 1995;76(7):1197–1200.
6. Gauthier ML, Berman HK, Miller C, Kozakeiwicz K, Chew K, Moore D, Rabban J, Chen YY, Kerlikowske K, Tlsty TD. Abrogated response to cellular stress identifies DCIS associated with subsequent tumor events and defines basal-like breast tumors. *Cancer Cell.* 2007;12(5):479–491.
7. Hu M, Peluffo G, Chen H, Gelman R, Schnitt S, Polyak K. Role of COX-2 in epithelial-stromal cell interactions and progression of ductal carcinoma in situ of the breast. *Proc Natl Acad Sci USA.* 2009;106(9):3372–3377.
8. Sathirapongsasuti JF1, Lee H, Horst BA, Brunner G, Cochran AJ, Binder S, Quackenbush J, Nelson SF. Exome sequencing-based copy-number variation and loss

of heterozygosity detection: ExomeCNV. Bioinformatics. 2011;27(19):2648-54.
9. Hart SN1, Therneau TM, Zhang Y, Poland GA, Kocher JP., Calculating sample size estimates for RNA sequencing data. J Comput Biol. 2013 Dec;20(12):970-8. doi: 10.1089/cmb.2012.0283.
10. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012), Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012 Mar 1;7(3):562-78. doi: 10.1038/nprot.2012.016.

From: eirb@mc.duke.edu
To: Amy Hobeika, Ph.D.
Subject: eIRB: Continuing review approved
Date: Monday, February 15, 2016 12:29:02 PM

IRB NOTIFICATION OF CONTINUING REVIEW APPROVAL

Continuing Review ID: CR001_Pro00059726
Principal Investigator: Herbert Lyerly
Protocol Title: Molecular Framework of Early Breast Cancer
Duke Cancer Center
Sponsor/Funding Source(s): US Department of Defense

Federal Funding Agency ID: W81XWH-14-0111
Date of Declared Concordance with federally funded grant, if applicable: N/A

The Duke University Health System Institutional Review Board for Clinical Investigations has conducted the following activity on the study cited above:

Activity: Continuing Review **Review Type:** Expedited
Review Date: 2/15/2016
Issue Date: 2/15/2016
Anniversary Date: 3/9/2016
Expiration Date: 3/9/2017

DUHS IRB approval encompasses the following specific components of the study:

Protocol, version/date:	--
Summary, version/date:	--3/5/2015
Consent form reference date:	--
Investigator Brochure, version/date:	--
Pediatric Risk Category:	--
Other:	--

The DUHS IRB has determined the specific components above to be in compliance with all applicable Health Insurance Portability and Accountability Act ("HIPAA") regulations.

This study expires at 12 AM on the Expiration Date cited above. At that time, all study activity must cease. If you wish to continue specific study activities directly related to subject safety, you must immediately email Jody Power at jody.power@duke.edu or call the IRB Office at 668-5111 and follow the instructions to reach the IRB Chair on call. Continuing review submissions (renewals) must be received by the DUHS IRB office 60 to 45 days prior to the Expiration Date.

No change to the protocol, consent form or other approved document may be implemented without first obtaining IRB approval for the change. Any proposed change must be submitted as an amendment. If necessary in a life-threatening situation, where time does not permit your prior consultation with the IRB, you may act contrary to the protocol if the action is in the best interest of the subject. You must notify the IRB of your action within five (5) working days of the event.

The Duke University Health System Institutional Review Board for Clinical Investigations (DUHS IRB), is duly constituted, fulfilling all requirements for diversity, and has written procedures for initial and continuing review of human research protocols. The DUHS IRB complies with all U.S. regulatory requirements related to the protection of human research participants. Specifically, the DUHS IRB complies with 45CFR46, 21CFR50, 21CFR56, 21CFR312, 21CFR812, and 45CFR164.508-514. In addition, the DUHS IRB complies with the Guidelines of the International Conference on Harmonization to the extent required by the U. S. Food and Drug Administration.

DUHS Institutional Review Board
2424 Erwin Rd | Suite 405 | Durham, NC | 919.668.5111
Federalwide Assurance No: FWA 00009025