

Name	Jessica Lal
Institution	The University of Texas at Arlington
Address	665 W. Mitchell St. Arlington, TX, 76005 United States
Phone Number	817-272-3569
Alternate Phone Number	832-788-5909
Email Address	jessica.lal@uta.edu

Project Requirements and Description

Project Title	Network-Based Genomic Determinants of Anthracycline Cardiotoxicity in Childhood Cancer Survivors
----------------------	--

Planned research population (eligibility criteria)

Anthracycline-exposed childhood cancer survivors enrolled in CCSS and St. Jude LIFE with available germline genomic data, treatment-exposure information, and cardiac outcome data. CCSS will serve as the discovery cohort and St. Jude LIFE as an independent replication cohort, excluding overlapping participants. Secondary analyses will evaluate sarcoma and AML survivors separately.

Proposed specific aims

The overall objective of this study is to determine how germline genomic variation within protein–protein interaction networks relate to susceptibility to cancer treatment–related cardiomyopathy (CTRCD) in childhood cancer survivors. We will test whether network-derived genomic features perturb functionally connected protein-protein networks and improve long-term cardiomyopathy risk prediction. Our central hypothesis is that susceptibility is driven in part by coordinated perturbation of functionally connected protein networks and that network-derived genomic features provide predictive information beyond treatment exposure and clinical risk factors. CCSS will serve as discovery and St. Jude LIFE as an independent replication, using complementary cardiac phenotypes and excluding overlapping participants from replication. Analyses using CCSS and St Jude LIFE data will be conducted in collaboration with Dr. Yadav Sapkota in the Department of Epidemiology and Cancer Control.

Specific Aim 1: Identify genomic susceptibility features associated with anthracycline cardiotoxicity using outcome-informed representation learning.

We hypothesize that germline genomic variation contains latent, nonlinear patterns associated with susceptibility to CTRCD. Using CCSS, we will develop an outcome-informed autoencoder with separate training and held-out test sets, performing cross-validation within the training set for feature selection and model tuning. The model will encode high-dimensional genomic loci into a lower-dimensional latent space while incorporating cardiac outcomes and relevant clinical and treatment characteristics. We will interpret cardiotoxicity-associated latent representations using SHAP-based feature attribution to identify contributing genomic loci and enriched biological pathways. We will assess stability across cross-validation folds in CCSS and evaluate predictive performance in the held-out test set. Expected outcome: Identification of stable, biologically interpretable latent genomic susceptibility to CTRCD and prioritized loci for network analysis. Specific Aim 2: Define protein-network modules underlying genomic susceptibility to anthracycline cardiotoxicity.

We hypothesize that susceptibility loci identified in Aim 1 converge within functionally connected protein network modules representing biological pathways involved in CTRCD. We will map prioritized loci to

proteins in a curated human protein–protein interaction network and quantify proximity to disease-associated proteins using shortest-path distances, testing significance against 10,000 random protein sets matched for size and degree. Secondary CCSS analyses will examine sarcoma and acute myeloid leukemia (AML) survivors separately. Expected outcome: Identification of biologically coherent protein-network modules linking inherited genomic variation to CTRCD.

Specific Aim 3: Determine whether network-derived genomic features improve prediction of anthracycline cardiotoxicity.

We hypothesize that network-derived susceptibility features provide complementary predictive information beyond established clinical, treatment, and genomic risk factors. We will develop models integrating network features with demographic, treatment, and established genomic risk factors, including GWAS loci and polygenic risk scores, using cross-validation in CCSS, and evaluate their performance in St. Jude LIFE as an unseen, independent replication cohort. We will evaluate model performance using ROC AUC, sensitivity, specificity, PPV, and NPV, each with 95% confidence intervals for each measure. Expected outcome: Determination of whether network-derived features provide reproducible improvements in long-term CTRCD risk prediction.

Impact: This work will establish a network-medicine framework linking germline genomic variation to biological susceptibility and long-term cardiac outcomes, providing a foundation for improved risk stratification and future multi-omic studies of anthracycline cardiotoxicity

Will the project require non-CCSS funding to complete?

Yes

If yes, what would be the anticipated source(s) and timeline(s) for securing funding?

Analyses are expected to support ongoing and planned externally funded research focused on molecular susceptibility to treatment-related late effects, including an NIH R01 application. Institutional start-up resources are also available to support pilot molecular analyses. Additional external funding will be sought for laboratory-based assays or ancillary biospecimen studies that extend beyond existing CCSS/SJLIFE data resources.

Does this project require contact of CCSS study subjects for:

Additional self-reported information	No
Biological samples	No
Medical record data	No

If yes to any of the above, please briefly describe.

What CCSS Working Group(s) would likely be involved? (Select all that apply)

	Primary	Secondary
Second Malignancy		
Chronic Disease		✓
Psychology/Neuropsychology		
Genetics	✓	
Cancer Control		

Epidemiology/Biostatistics		✓
----------------------------	--	---

Outcomes or Correlative Factors

	Primary	Secondary	Correlative Factors
Late Mortality		✓	
Second Malignancy			

Health Behaviors

	Primary	Secondary	Correlative Factors
Tobacco			✓
Alcohol			✓
Physical Activity			✓
Medical Screening			✓
Other			✓

If other, please specify

Psychosocial

	Primary	Secondary	Correlative Factors
Insurance			✓
Marriage			✓
Education			✓
Employment			
Other			

If other, please specify

Medical Conditions

	Primary	Secondary	Correlative Factors
Hearing/Vision/Speech			
Hormonal Systems			
Heart and Vascular	✓		
Respiratory			
Digestive			
Surgical Procedures			
Brain and Nervous System			
Other			✓

If other, please specify

we will correct our statistical models for patients with history of type II diabetes, obesity, chronic kidney disease, hyperlipidemia, high blood pressure

Medications

Describe medications

Anthracyclines

Psychologic/Quality of Life

	Primary	Secondary	Correlative Factors
BSI-18			
	Primary	Secondary	Correlative Factors
SF-36			
CCSS-NCQ			
PTS			
PTG			
Other			

If other, please specify

Other

	Primary	Secondary	Correlative Factors
Pregnancy and Offspring			
Family History			
Chronic Conditions (CTCAE v3)		✓	
Health Status			

Demographic

	Primary	Secondary	Correlative Factors
Age			✓
Race			✓
Sex			✓
Other			

If other, please specify

Cancer Treatment

	Correlative Factors
Chemotherapy	✓
Radiation Therapy	✓
Surgery	

Anticipated Sources of Statistical Support

CCSS Statistical Center	No
Local Institutional Statistician	No

If local, please provide the name(s) and contact information of the statistician(s) to be involved.

Will this project utilize CCSS biologic samples? No

If yes, which of the following?

If other, please explain

Other General Comments

This is a revised submission of the previously proposed study, “Molecular Susceptibility Endophenotypes of Treatment-Related Late Effects in Childhood Cancer Survivors.” The revised proposal has been narrowed to focus specifically on germline genomic susceptibility to anthracycline-associated cardiomyopathy and the identification of network-defined genomic features that may explain heterogeneous cardiac outcomes among similarly exposed survivors. It is intended to complement, rather than duplicate, ongoing CCSS investigations of individual genetic predictors of cardiomyopathy by emphasizing integration of germline variation within protein–protein interaction networks and evaluation of network-derived susceptibility features in independent cohorts. Analyses will be conducted in collaboration with Dr. Yadav Sapkota in the Department of Epidemiology and Cancer Control.
