University of Pennsylvania

Annual Progress Report: 2005 Formula Grant

Reporting Period

July 1, 2009 – December 31, 2009

Formula Grant Overview

The University of Pennsylvania received $9,229,640 in formula funds for the grant award period January 1, 2006 through December 31, 2009. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Effects of Nicotine Deprivation and COMT Genotype on Cognitive Performance - Identify the neural and genetic substrates of cognitive deficits experienced during nicotine deprivation in dependent smokers.

Duration of Project

1/1/2006 – 9/1/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement C.U.R.E. Annual Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 2: Project Title and Purpose

Developing Effective Gene Silencing Agents for Human Malignancies - Advance the development of messenger RNA (mRNA) "silencing" molecules for the treatment of hematologic malignancies. When an mRNA is silenced the information it contains can no longer be used to instruct the cell's machinery to make the specific protein it encodes. If the proper mRNA targets are selected a cell's ability to grow, or survive, could be greatly compromised. Accordingly, this strategy should be very effective in the treatment of many different types of malignancies.

Duration of Project

1/1/2006 - 12/31/2009

Project Overview

The goal of this project is to advance the development of nucleic acid based therapeutics for the treatment of hematologic, and eventually other, malignancies. Its fundamental hypothesis is that appropriately targeted, and delivered, antisense nucleic acids (ASNA), including oligodeoxynucleotides (ODN) and short interfering/short hairpin RNA (siRNA/shRNA), can effectively control malignant cell growth, or kill such cells outright with little damage to normal tissues. We further hypothesize that this treatment strategy can obviate the evolution of "drug resistant" cells, as has been observed with small molecule drugs such as imatinib, because if synthesis of a required protein is completely suppressed, point mutations cannot be selected for. This project will focus on novel chemical backbones modifications and development of scale up delivery methods so that these materials can be used to develop purging protocols for lymphomas and myeloma. It has 2 specific aims, which are as follows:

Aim #1: Develop chemically modified nucleic acid drugs that will efficiently silence candidate mRNA targets.
Aim #2. Examine the ability of rationally designed antisense nucleic acids to kill primary human cancer cells in immunocompromised mice.

Aim #1 will test a variety of novel backbone chemistries and delivery methods. These, combined with rational targeting, should lead to development of candidate antisense molecules with enhanced gene modulating activity. Aim #2 will test this assumption in a xenogenic transplant model employing wild type leukemia cells obtained from the tissue bank core in a NOD/SCID animal model system. Uptake, distribution, and clearance of the molecules from various fluid and organ compartments will be studied using a variety of hybridization and chromatographic techniques. Data will be analyzed as a function of time, dose, and schedule of drug administration in order to facilitate development of rational drug delivery and treatment protocols.

Principal Investigator

Alan M. Gewirtz, MD
University of Pennsylvania
School of Medicine
Rm 713, BRB II/III
Dept. of Medicine
421 Curie Blvd.
Philadelphia, PA 19104

Other Participating Researchers

Yuji Nakata, MD - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

The outcome of this research should be more effective and clinically useful gene silencing reagents. At the present time, molecules available for this purpose have not been particularly effective in the clinic. We postulate that this is because the molecules employed for these studies were not efficiently delivered or failed to bind to their target. The investigations to be employed during the conduct of this research project will address these issues, and in so doing, lead directly to the development of highly effective gene silencing agents for the treatment of cancer and other diseases.

Summary of Research Completed

We remain keenly interested in silencing the c-myb protooncogene because of its critical role in regulating hematopoietic and immune cell development, and the high likelihood that abrogating c-myb function will be of use in treating hematologic, breast and colon malignancies, as well as autoimmune disease. These translational goals will likely require that longer term suppression of c-myb function will need to be obtained. For this reason, during this reporting period we re-visited the use of short hairpin RNAs (shRNA) for longer term silencing of c-myb mRNA. While carrying out these studies, we also identified a new function for c-myb, which provided mechanistic details as to how c-myb regulates the switch between Th1 and Th2 cel development. These studies have been submitted to Blood, and a revision has been requested so we are hopeful that they will soon be published. An abstract of this work follows:

The c-Myb and GATA-3 transcription factors play important roles in T-cell development but the mechanistic details of how they interact, especially during human peripheral blood T-cell development, are incomplete. Herein, we report that GATA-3 and c-Myb form a transcription complex essential for development and maintenance of primary human Th2 cells. Silencing c-Myb in normal human naive CD4+ cells under Th2 cell promoting conditions blocked up-regulation of GATA-3 and IL-4, and in effector/memory CD4+ T-cells, decreased expression of GATA-3 and Th2 cytokines. We also found that in primary T-cells, c-Myb allows GATA-3 to auto-activate its own expression, an event that requires the direct interaction of c-Myb and GATA-3 on their respective binding sites in GATA-3's promoter. Interestingly, when immunoprecipitated, the c-Myb/GATA-3 complex contains Menin, a tumor suppressor protein that targets c-Myb to MLL. In developing Th2 cells and in CD4+ effector/memory cells, c-Myb, GATA-3 and Menin bound to the same region in the GATA-3 promoter. Recruitment of MLL into the c-Myb, GATA-3, Menin complex was associated with the formation Th2 memory cells. Therefore, c-Myb, GATA-3, and Menin form a core transcription complex that regulates GATA-3 expression and ultimately Th2 cell development in primary peripheral blood T-cells.

<u>Lentivirus vector construct for expressing c-myb shRNA</u>
c-Myb-1(tgttattgccaagcactta), c-myb-3 (ctgcctggacgaactgata) and control c-myb-1 scramble (ctttatacgtagtcataag) siRNA sequences were inserted into the H1UG lentivirus vector (a gift of Dr. E.J. Brown, University of Pennsylvania) which was modified from FUGW vector. The

mixed lentivirus constructs expressing c-myb-1 and c-myb-3 shRNA were transduced into human primary CD4+ T-cells.

Transduction of Lentivirus to human primary CD4+T-cells

Human primary CD4+ T cells were stimulated with anti-CD3/CD28 antibodies (10 µL/2-5x10^6 cells of Dynabeads CD3/CD28 T cell expander, Invitrogen) for 16 hours, followed by IL-2 (10ng/ml) for 5 hours in RPMI + 10% FCS. Lentivirus expressing GFP + c-myb-1, and c-myb-3, or scrambled sequence control shRNA were freshly prepared in 293FT cells (Invitrogen). The Lentivirus were added to the stimulated CD4+ T-cells (CD4+ T-cells 5x10^6/ml:viral supernatant = 1:1.5 vol/vol, Polybrene 8 mg/ml). Five days post-transduction, GFP positive cells were sorted by the FACSAria™ cell sorter (BD Biosciences). Transduction efficiency ranged from 10-50 % of live cells (Figure 1). The GFP positive cells were further sorted using magnetic beads conjugated with anti-CD45RO+ antibody (CD45 MicroBeads, Mileni Biotec). Alternatively, GFP+ CD4+ T-cells were sorted with CD45RA-PE antibody using the FACSAria™ cell sorter. The cells selected negatively with CD45RO, or positively CD45RA positive cells, were classified as naïve CD4+ T-cells.

Figure 1. Typical flow analysis of shRNA transduced cells. The live cells were 46.8 % of all events, and GFP+ cells were 41.3% of live cells (19.0% of all cells).

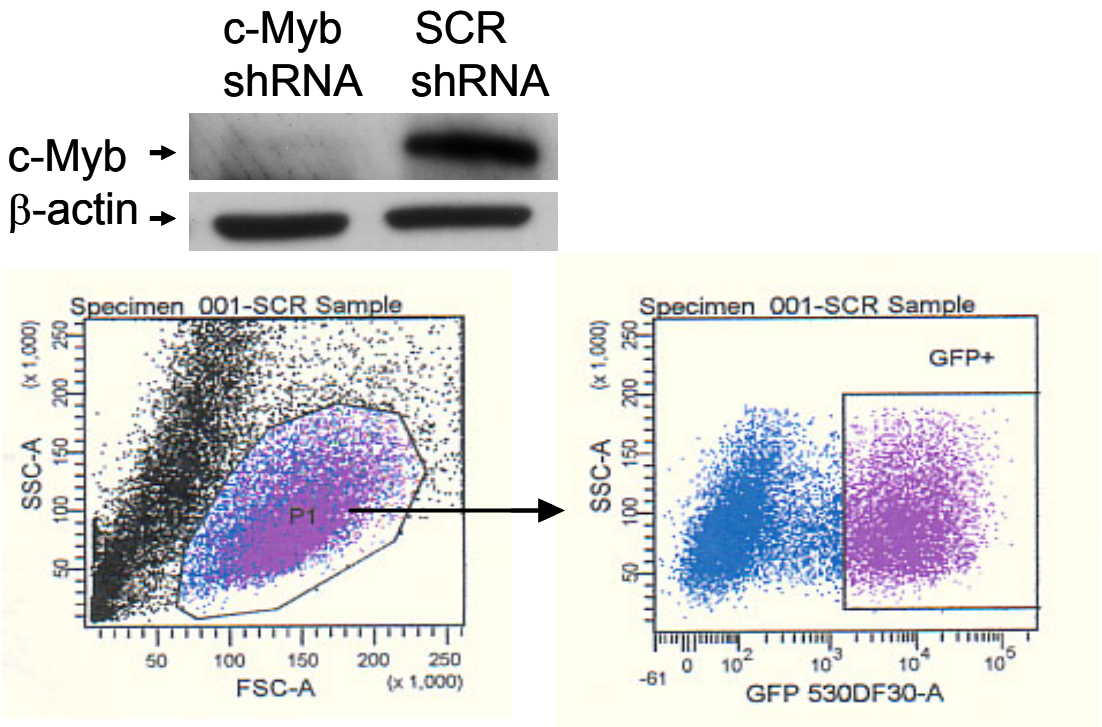


Figure 2 c-Myb protein suppression by c-Myb shRNA Lentivirus system in the naïve T-cells under Th2 promoting conditions. Normal human peripheral blood CD4+ cells were transduced with lentivirus constructs expressing GFP + c-myb shRNA. Sorted GFP positive CD4+CD45RO- cells were cultured under Th2 promoting conditions for 5 days. Western blotting was carried out with anti-c-Myb antibody. SCR= scrambled shRNA.

Research Project 3: Project Title and Purpose

Genetic Determinants of Bone Loss and Ovarian Failure in Breast Cancer Survivors - Identify genetic variants in reproductive hormone and vitamin D pathways that are associated with premature ovarian failure and/or bone loss in survivors of breast cancer who received adjuvant chemotherapy. By identifying patients at risk for these late effects of therapy, interventions will be developed to prevent or ameliorate these potentially devastating long term effects of curative breast cancer therapies, decrease symptoms that adversely impact quality of life and prevent long-term skeletal and other complications that reduce function and productivity in this population.

Duration of Project

1/1/2006 – 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Infrastructure Project 4: Project Title and Purpose

Research Infrastructure to Establish a State-of-the-Art Tumor Tissue and Biospecimen Bank - To establish a state-of-the-art Tumor Tissue and Biospecimen Bank to provide immediate and long-term access to properly processed and maintained malignant and normal tissue for current and future cancer research. The infrastructure funds will allow renovation of dedicated space that will meet the requirements of the facility to properly collect, process, monitor, and store (long and short term) tissue samples and blood components.

Duration of Project

1/1/2006 - 12/31/2009

Project Overview

The objective of this research infrastructure project is to establish a state-of-the-art Tumor Tissue and Biospecimen Bank that allows easy access to properly collected and maintained normal and malignant tissue and blood components that are needed to conduct competitive clinical research now and, as new technologies are developed, well into the future.

This objective will be met by renovation of an identified, dedicated and specialized facility with properly designed equipment meeting the highest standards of air quality to achieve the following specific aims:

- Collection and storage of normal and malignant tissue and blood components that meets or exceeds standards NCI Best Practices for Biorepositories (June 2007) and the ISBER Best Practices (March 2005)
- Identification and accurate and complete pathological descriptions of all tissues and blood components by clinically trained personnel.
- Establish a bioinformatics structure for rapid response to an investigator's request, sample identification and correlation of associated data including results from previous testing (e.g., microarray analysis) and/or research findings.

Principal Investigator

Michael Feldman, MD, PhD
Abramson Cancer Center
University of Pennsylvania
3400 Spruce Street
Philadelphia, PA 19104

Other Participating Researchers

Mariusz A. Wasik, MD, Lisa B. Miranda, BS, CCRC - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

The establishment of the Tumor Tissue and Biospecimen Bank is expected to ultimately result in the identification of new treatments for cancer. Immediate benefits include:

- Allow investigators to conduct innovative and competitive research due to the ease of access to normal and malignant tissue that is properly prepared and stored. Adhering to Best practices & good clinical practice guidelines will also increase the reliability of the data collected.
- Increase the number of samples available for research due to the ability to access results from previously performed tests and experiments so as not to duplicate efforts.
- Identify factors giving rise to different tumor types or factors involved in health disparities for different groups or populations.

Summary of Research Completed

The infrastructure research funds this last year were spent to establish a state-of-the-art Tumor Tissue and Biospecimen Bank that allows easy access to properly collected and maintained normal and malignant tissue and blood components that are needed to conduct competitive clinical research now and, as new technologies are developed, well into the future. The infrastructure funds helped to begin the establishment of the three areas as described:

- Collection and storage of normal and malignant tissue and blood components that meets or exceeds standards NCI Best Practices for Biorepositories (June 2007) and the ISBER Best Practices (March 2005)
- Identification and accurate and complete pathological descriptions of all tissues and blood components by clinically trained personnel.

- Establish a bioinformatics structure for rapid response to an investigator's request, sample identification and correlation of associated data including results from previous testing (e.g., microarray analysis) and/or research findings.

The basement of the Maloney Building was completely renovated for the Tumor Tissue and Biospecimen Bank (TTAB). Included in the renovation was a new heating and ventilation system (HVAC) that meets the NCI Best Practices for Biorepositories and ISBER Best Practices. Emergency Power Back-up was also installed and was necessary for the protection against loss of samples. The infrastructure of the bank includes high capacity freezers and allows for redundancy in the system. A new freezer monitoring system was also installed that is internet based and meets the 21CFR11 regulation for compliance.

Collection and storage of tumor tissues and biosamples

Since its inception, the tumor tissue and biosample bank (TTAB) has sought to collect paired sampled of tumor and normal tissue along with associated fluid biosamples using NCI best practice guidelines. In this past funding year, funds were used for partial salary support for setup and configuration of the fully operational bank, including configuring freezers, freezer monitoring system, developing and documenting standard SOP's, manage banking requests and provide histophathology services to investigators. The SOP's that have been developed are shared with other banking efforts at the School of Medicine in a school wide intitiatvie to coordinate standardized biosample banking efforts across the school. Together, these efforts resulted in the collection of 6347 biosamples from cancer patients at the University of Pennsylvania Health System. These collections supported cancer researchers in multiple areas of investigation including the following (Table 1). These samples supported tissue collections across 18 different organ specific tumor groups.

Table 1: Biosamples by study area

Study group	Number samples
Multiple Myeloma	387
Non-group related (broad range of cancers)	2495
Gyn Oncology group	1204
Breast Cancer group	3
Renal cell carcinoma group	27
Viral tumor group	135
Warfarin Anticoagulation	899
Prostate Cancer group	162
Mesothelioma group	1035

All of these samples are stored and managed in our tissue inventory application, caTissue which is part of the NCI's caBig program. By using caTissue, the samples in our bank are searchable on the NCI's cancer bioinformatics grid. To accomplish this, a grid node was established for biosample query from other non-Penn NCI caBig so that scientists can search for samples in a de-identified manner. By exposing our biosample repository on the NCI's bioinformatics grid, we are fostering cross institutional collaborations. To date, we sent material to 4 outside investigators using samples from our mesothelioma group collections. We also received requests

for breast and melanoma samples. The melanoma request was sent to our Skin Spore group. The request for breast cancer samples could not be filled as we are still in the early developmental stages of breast cancer banking program with our breast research group.

As part of these tissue banking operations, samples are routinely QC'ed by Dr. Feldman. This QC is recorded based on pathologic review of the banked samples and is evaluated for tumor type, amount of tumor in sample and percent tumor necrosis. All of these variable are collected on every banked sample and the QC review data is entered into caTissue.

TTAB's banking efforts have allowed Penn's investigators to grow their research efforts and to procure new grant funding. Our myeloma group successfully competed for a K23 award for one of it's junior investigators. Our Ovarian Oncology Center of Excellence successfully renewed their ovary spore grant with the NCI. As part of that Ovarian Spore grant, TTAB was identified as the principle tissue banking team for Penn's Ovarian Cancer Researchers. Penn's Renal Cell Carcinoma group obtained an RO1 to study tumor response to Sorafenib and Sumitenib in clear cell renal cell carcinoma (TTAB will be performing all banking operations for this grant).

BioInformatics Structure to support TTAB

This year funds were used to provide software customization costs to enhance caTissue. The software customizations were used to extend caTissue so that biosamples that are tracked and managed in TTAB freezers can be tracked and managed as they are moved to our molecular genotyping core facility for nucleic acid extraction and analysis. The resulting molecular samples and their associated QA/QC metrics (OD260, OD280, Sample amounts, URL to gel trace for sample quality, sample types, derivatized sample ID's and assays performed on the molecular samples) can all be collected and loaded into caTissue using excel spreadsheets with automated uploads and data integrity checking. This need for sample tracking from the bank through common analytic laboratories was identified as a critical shortcoming of caTissue v1.1 which necessitated software enhancements. The software changes were implemented in collaboration with our end users and the molecular analysis core laboratory.

This software solution is now being used by our Warfarin study group and will also be used in 2010 by our tobacco cessation study group and lung carcinoma group. The specific use cases for this software enhancement were developed in collaboration with our smoking cessation study group.

Using caTissue, Penn has the largest collection of biosamples exposed on the NCI's caBIG bioinformatics grid. Our 16,000 samples are available for query by scientists across any NCI cancer center in the USA.

Advanced tissue analytic workstations

In addition to biosample banking, investigators have been requesting tissue analytic services from TTAB. Dr. Feldman has been developing an microscopic tissue analytic system for measuring pathologic tissue stained with multiple antibodies (in bright field or fluorescence) and analyzing the tissue with advanced computer algorithm's designed to measure antibody's at the cellular and subcellular level both within tumor compartments as well as non-tumor compartments. This analytic platform, which we have termed Histocytometry, is equivalent to

flow cytometry on fixed paraffin samples on glass slides. The resulting cellular data is expressed as a list mode data file which shows on a per cell basis the amount of each antibody in the nucleus, cytoplasm or membrane areas for either tumor cells or non-tumor cells. This advanced analytic approach is now being used for multiplexed immunochemical studies of tumors in their native environment, yielding new insights into the biologic heterogeneity that is present in tumors that may be lost by destructive analytic processes like western blot analysis. Money from this year's Tobacco grant was used to purchase two analytic workstations to run the software algorithmns so that the computer that runs the robotic multispectral image capture system could be used for image capture and not tied up by users for data analysis. The imaging core of the TTAB is now fully functional with a robotic image capture system to study multiplexed antibody stains in tissues or cell preparations. The two new analytic workstation round out the Histocytometry suite along with 8 Terabytes of shared network storage. This new Histocytometry lab is being used by multiple investigators including both Spore groups (Ovarian and Skin), the renal cell carcinoma group, the prostate carcinoma group and soon, the breast carcinoma group.

Research Infrastructure Project 5: Project Title and Purpose

Developing a Stem Cell Research Facility at the Penn School of Veterinary Medicine - The Department of Animal Biology at the Penn School of Veterinary Medicine (SVM) wishes to build vigorous stem cell research as part of a proposed University wide Center for Stem Cell Biology and Regenerative Medicine. Our objective is to stimulate research on the treatment of neurodegenerative disease (Alzheimer's, Parkinson's), dilated cardiomyopathy, osteoporosis, diabetes, cancer, hematopoietic and reproductive disorders by consolidating our existing strength in this area and by the recruitment of additional faculty. The purpose of this application is to develop research infrastructure by renovating about 4400 square feet of space into a modern research laboratory and core facility for stem cell research. Funds will be used for outfitting 4400 square feet of space in the Rosenthal Building of SVM with a modern research laboratory and stem cell core facility for cell derivation, maintenance, and cryopreservation.

Duration of Project

7/1/2007 - 12/31/2009

Project Overview

The School of Veterinary Medicine (SVM), at the University of Pennsylvania, is recognized as a world leader in stem cell biology and the birth place of mouse transgenesis. Stem cell studies will be pervasive and critical throughout biology and medicine in the coming years, and the use of stem cells as therapeutic agents presents an enormous potential for treating a wide spectrum of degenerative disease and cancer. SVM proposes to establish a strong stem cell component in the Department of Animal Biology, particularly focused on the use of adult stem cells with hope that this would provide impetus for starting a major Stem Cell Center at the University. A volume of basic research has indicated an exciting possibility of using stem cell graft or stem cell replacement therapy for treating a number of dreadful genetic and degenerative diseases such as Alzheimer's, Parkinson's, breast and hematopeitic cancers, and bone degenerative diseases such

as osteoporosis. As part of the new University-wide Center, SVM proposes to build a vigorous stem research program for developing treatment modalities for these diseases using animal models that are either presently available or patients that are presented to the Ryan Small Animal Hospital at Penn.

A specific objective of this current project is to renovate about 4400 square feet of space in the Rosenthal Building located at the 38th and Spruce Street into much needed laboratory space and core facility. Part of the funds would be used to enhance SVM's laboratory offering, both in terms of space and equipment, to house stem cell researchers. The spatial needs of a Stem Cell research are extensive and SVM plans to overcome this problem by outfitting and equipping about 3400 square feet of the available space in the Rosenthal building. This includes, but is not limited to, bench space for researchers, equipment space, surgical facility, molecular biology facility, and office space. The remaining 1000 square feet of space will be renovated into a Stem Cell Biology Core facility which is critical for generating adult cardiac, bone (mesenchymal), brain and embryonic stem cells. The latter will contain facility for stem cell derivation, BSL2 facility for culturing and maintenance of stem cells, and crypreservation.
We envision that the proposed research project will stimulate research collaboration and interaction between basic stem cell researchers and clinical scientists for exploiting the wealth of animal models presented to our veterinary clinic and developing stem cell based treatment modalities.

Principal Investigator

Steven Fluharty, PhD
Associate Vice Provost for Research
University of Pennsylvania
Office of the Vice Provost for Research
118 College Hall
Philadelphia, PA 19104

Other Participating Researchers

Narayan Avadhani, PhD, Alan Kelly, PhD, Ralph Brinster, V.MD, PhD, Richard Schultz, PhD, Steven Emerson, MD, PhD - employed by University of Pennsylvania
Richard Behringer, PhD - employed by MD Anderson Cancer Center
Kurt Hankenson, DVM., PhD - employed by University of Michigan
Hans Scholer, PhD - employed by Max Plank Institute for Stem Cell Biology, Munster, Germany

Expected Research Outcomes and Benefits

The proposed Center for Stem Cell Research will not only help consolidate the School of Veterinary Medicine's (SVM) leadership position in basic biology of stem cells, but will also help stimulate research on the direct application for treating human and animal diseases. This research infrastructure project should help Penn assume a leadership role in developing stem cell based therapies for the treatment of various degenerative diseases and cancer in addition to developing methods to introduce new genes in whole animals through stem cells for modulating

the immune system. This could provide a powerful approach to combat cancer and diseases of the immune system. Specifically, the outcome of this project is likely to include the use of Mesenchymal Stem Cells (MSC), a stem cell of the bone marrow for inducing bone regeneration in animal models and human patients with osteoporosis and devastating athletic injuries; use of hematopeitic stem cells for treating various cancers; use of adult brain stem cell transplantation for treating Alzheimer's disease and correcting myocardial dysfunction by stem cell transplantation. This project is also expected to significantly enhance the academic environment for researchers and students by stimulating interaction of stem cell biologists with members of SVM's Center for Medical Genetics, Mari Lowe Center for Comparative Oncology and Center for Infectious Disease, as well as a number of Centers at the Medical School. Location of SVM and the Department of Animal Biology in the center of a major biomedical research campus including Penn's School of Medicine, Dental Medicine, School of Arts & Sciences, the Children's Hospital, and the Wistar Institute presents an unparalleled advantage of conducting multidisciplinary research in animal models with potential usefulness in treating human diseases.

Summary of Research Completed

Project funds were expended on construction of the laboratory. The School consolidated the plans for the location of the Center for Stem Cell Research and completed the construction of the laboratory and associated office space. Primary Investigators are planning their move into the facility, which should be partially occupied by late spring 2010.

Facilities

A long-standing plan to build Penn Vet's regenerative medicine program has come to fruition with the completion of renovations on the third floor of Rosenthal Building. The approximately 6000 square foot space includes:

5 tissue culture rooms
5 flexible procedure rooms
2 equipment storage areas
1 conference room
45 work stations for students/technicians
295 linear feet of bench space
1 cold room
5 new faculty offices

The space, greatly improved by the renovations, provides research laboratories that will eventually house five faculty members and their research teams.

Faculty

Currently, two faculty members, Drs Frank Golder and Ralph Meyer, have moved into the new laboratories. Penn Vet's faculty expertise in germ and stem cell biology is supported and facilitated by the new space and it will accommodate additional positions to be recruited in the coming years. In the emerging initiative of regenerative medicine, the School foresees recruitment of two new faculty members with research interests in 1) neural stem cells,

neurodegenerative disease, 2) the role of mesenchymal stem cells in treating osteoporosis, and 3) transdifferentiation of pluripotent stem cells in the treatment of diabetes.

Research Interests of the Faculty Located in the Newly Renovated 3rd Floor of Rosenthal Building

Francis John Golder, B.V.Sc., Ph.D.
Diplomate of the American College of Veterinary Anesthesiology
Assistant Professor in Anesthesiology, Department of Clinical Studies
School of Veterinary Medicine, University of Pennsylvania

Dr. Golder's long-term goal is to determine the cellular, neurochemical, and molecular basis of the functional plasticity of phrenic motoneurons that innervate the diaphragm. Long-term changes in motoneuron function can occur via plasticity of premotor synaptic connections. Using models of synaptic plasticity, he attributes changes in synaptic strength to neuromodulators - such as serotonin, the tropomyosin-related kinase B (TrkB) receptor ligands brain-derived neurotrophic factor (BDNF) and neurotrophin-4 (NT-4), and adenosine. His recent studies have revealed an exciting and novel model of respiratory motor plasticity that involves strengthening of spinal synaptic connections to phrenic motoneurons via spinal adenosine A2a receptor activation. A2a receptor-induced phrenic motor facilitation increases tidal volume in spinally injured rats. Thus, these studies may have important implications for recovery of respiratory function in human and veterinary patients experiencing respiratory insufficiency after spinal cord injuries and during motoneuron degenerative disease. His present research is focused on determining the cellular mechanisms contributing to adenosine A2a-induced phrenic motor facilitation.

Ralph Guenter Meyer, PhD
Assistant Professor, Department of Animal Biology
School of Veterinary Medicine, University of Pennsylvania

Dr. Meyer's research focuses on the role of poly ADP-ribosyl polymerase associated DNA repair in meiosis, which is a critical chromosome segregation stage in spermatogonial stem cell differentiation. The notion that DNA breakage and rejoining occurs in stem cell differentiation is a new concept, and Dr. Meyer is interested in how poly ADP ribosylation plays a critical role in this process.

Research Project 6: Project Title and Purpose

Outcomes of an Exercise Program for Older African American Women in a PACE Model - In this project we will learn whether a 16 week structured exercise program given three times a week for older African American women enrolled a program in Philadelphia called Living Independently For Elders, which provides services to avoid moving to a nursing home, will improve the participants' strength and endurance, help participants avoid becoming depressed, and improve their memory and thinking. We will also learn if blood pressure and weight, for those who are overweight, are lowered and sleep improves for those who participate. This will

help support funding for more community-based exercise programs which improve the health and well being of older African American women.

Duration of Project

1/1/2006 - 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Infrastructure Project 7: Project Title and Purpose

Vivarium – Research Infrastructure (Construction Phase II) - To breed and study specialized mice and other rodent models which have specific genes deleted, added or altered. Mouse models are our most important tool to study both diseases with a genetic component and to understand normal body processes which may be altered to result in disease. Mice are the ideal model because they are mammals and share many similarities in basic body functions to humans, while they are small, breed rapidly and have been a major focus of scientific study. This project is a multi-year project including scope development, design, contract awards and construction.

Duration of Project

1/1/2006 – 6/30/2006

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Infrastructure Project 8: Project Title and Purpose

Diagnostic Serology Laboratory for University Laboratory Animal Resources - Research Infrastructure - To improve turn-around time, optimize reliability, and optimize use of resources for routine and "outbreak" serologic testing which is essential to maintain the health of more than 70,000 experimental rodents used on our campus.

Duration of Project

1/1/2006 - 6/30/2006

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 9: Project Title and Purpose

Computational Cell Biology of Human Diseases - Many human diseases arise as a result of a dysfunction of cellular processes. For example, cancer is the result of dysregulation of cell growth and cell death. Similarly, certain neuropsychiatric diseases such as autism and schizophrenia can be traced to the loss of control of cell-to-cell communication processes in the brain. The purpose of this project is to develop new computational approaches and genome-scale measurement technologies to understand the molecular control of cell processes and how the loss of such control leads to human diseases.

Duration of Project

1/1/2006 - 12/31/2009

Project Overview

In the last ten years, tremendous advancements in genomics have brought forth biomedical data at an unprecedented scale and level of detail. These data have revealed to us that the molecular underpinnings of human diseases are unbelievably complex; that, a comprehensive understanding of the causal processes requires a completely different approach to integrating information and biological processes at different levels of organization. We will develop a novel quantitative computation-driven integrative research process, introducing a new organizational paradigm to biomedical research. Using this integrative research paradigm, we will strive to develop a comprehensive understanding of molecular control processes governing cellular function and genetic changes leading to loss of control and human diseases. More specifically, our aims are:

Aim 1: Develop a computational prediction of all genomic elements active in cellular processes. Computational prediction methods include using comparative analyses to detect novel functional elements, using comparative gene expression patterns to delineate regulatory motifs, using biophysics based models to confirm RNA structure, and predicting protein-protein interactions and protein-RNA interactions.

Aim 2: Develop tools for computational modeling of cellular processes and apply the model to genome scale data. Tools for computational modeling include integrated databases to collect basic information for modeling from distributed sources, natural language processing of biomedical literature to synthesize existing knowledge, dynamical systems models for generating predictions on cellular control processes.

Aim 3: Develop experimental procedures to measure cellular parameters for the computational model and survey sequence variation from patient populations to associate computational predictions to genetic variability underlying disease phenotypes.

Principal Investigator

Junhyong Kim, PhD
University of Pennsylvania
206 Goddard
Department of Biology
415 S. University Ave.
Philadelphia, PA 19104

Other Participating Researchers

Maja Bucan, PhD, Susan Davidson, PhD, James Eberwine, PhD, Vijay Kumar, PhD, Fernando Pereira, PhD, David Roos, PhD - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

Many human diseases such as autism, cancer, and infectious diseases can be traced to cellular processes and their dysfunction. Neuro-developmental disorders have a disproportionate impact on national health and society because they are often chronic and the available therapeutic approaches are meager. Infectious diseases affect every individual in the population. Finally, cancer has recently been reported have replaced heart diseases as the leading cause of death in US. And, because the cell is a compartmentalized system, these diseases are especially amenable to genome-scale data collection, in vivo measurements, and computational modeling. We expect to make significant advances in our understanding of how the disruption of cell control processes affects the normal function of specialized cells such as neurons. In particular, we expect to provide a quantitative process model for the molecular basis of human diseases. We expect to use our model of molecular control processes to associate genetic variation in disease populations, identifying the functional implication of genetic variation and enhancing the prediction of disease related genomic elements. The computational tools produced by this research program will also have wide applicability to other biomedical problems and the database produced by this program will be made available to the general biomedical community. We expect to develop novel in vivo measurement technologies for cellular processes, which will also have broad utility for genome-scale research of human diseases and molecular diagnostics.

Summary of Research Completed

In the last six months of this project, we focused on two key activities: finalizing publications of the research results from the project and organizing a team effort to seeking external funding for a systems biology center based on the results of the research.

In the final period, we prepared and submitted a seventh publication (Simola et al.). This paper describes new computational analysis of time-series data for cell cycles that suggested a new model of timing control in cell physiology.

Publications from the current project

Chen, Z., O. S. Harb and D. S. Roos (2008). "*In Silico* Identification of Specialized Secretory-Organelle Proteins in Apicomplexan Parasites and *In Vivo* Validation in *Toxoplasma gondii*." PLoS ONE **3**(10): e3611.

Guo, S. and J. Kim (2010). "Dissecting the molecular mechanism of *Drosophila* odorant receptors through activity modeling and comparative analysis." Proteins: Structure, Function, and Bioinformatics **78**(2): 381-399.

Kreher, S. A., D. Mathew, J. Kim and J. R. Carlson (2008). "Translation of Sensory Input into Behavioral Output via an Olfactory System." Neuron **59**(1): 110-124.

Lee, M. T. and J. Kim (2008). "Self Containment, a Property of Modular RNA Structures, Distinguishes microRNAs." PLoS Comput Biol **4**(8): e1000150.

Liu, Q., A. J. Mackey, D. S. Roos and F. C. N. Pereira (2008). "Evigan: a hidden variable model for integrating gene evidence for eukaryotic gene prediction." Bioinformatics **24**(5): 597-605.

Sul, J.-Y., C.-w. K. Wu, F. Zeng, J. Jochems, M. T. Lee, T. K. Kim, T. Peritz, P. Buckley, D. J. Cappelleri, M. Maronski, M. Kim, V. Kumar, D. Meaney, J. Kim and J. Eberwine (2009). "Transcriptome transfer produces a predictable cellular phenotype." Proceedings of the National Academy of Sciences **106**(18): 7624-7629.

As part of the project activities we organized a new team of interdisciplinary researchers (Table 1) to apply for a NIH Institute for General Medicine P50 Systems Biology Center grant. This application has been submitted (Oct, 2009) and is pending review. The abstract for the application is as follows:

Center for Synthetic Cell Programming: The main mission of the proposed Center for Synthetic Cell Programming is to develop a new top-down systems biology paradigm for synthetic biology and to develop educational and outreach programs centered on this paradigm. Synthetic Biology is a new discipline whose goal is to develop an engineering level understanding of biological systems to artificially construct new systems with pre-specified properties. A majority of current approaches in synthetic biology involve construction of modular components, such as particular gene regulation circuits, with the goal of scaling the components (i.e., bottom-up assembly) into larger, more complex systems. Here we propose an opposite approach of top-down engineering where systems are manipulated by targeting tunable parameters of the system; the parameters are adjusted until targeted properties are reached. In particular, we remodel the phenotype of mammalian cells using quantitative manipulations of entire transcriptomes (internal tunable variables), as well as inductive signals and durotactic stimulations (external tunable variables). We will leverage new technologies that we developed for laser-mediated cell transfection and microfabricated tissue engineering. In our preliminary investigations, we successfully used parameter tuning of (a) the transcriptomes to generate trans reprogramming of differentiated mammalian neurons and (b) the external durotactic cues to alter cellular phenotype properties.

The ultimate goal of our scientific program is to use state-switching systems biology principles to develop a top-down synthetic biology paradigm for reprogramming the phenotype of mammalian cells, an effort which is core to many efforts in biomedicine to both treat disease and understand key regulating steps in development.

The proposed center will bring together an interdisciplinary team of engineers, computational scientists, and biomedical researchers. The center will leverage the team's unique expertise in single cell manipulation, single cell molecular profiling, mechanobiology, and network analysis to achieve its goals. The center will broadly disseminate the techniques within the center to the biomedical community. In particular, the activities of the center are expected to produce novel instruments and components that can be disseminated directly and through open system plans. Leveraging the resources of the Penn Genome Frontiers Institute, the center will develop educational programs, especially at the faculty and the undergraduate level to promote systems biology to an underserved audience. The center will also develop outreach programs both to the general public and to high-school educators with the goal of emphasizing the role of engineering and technology in biomedical sciences. New faculty in systems biology will be recruited, especially with special attention to recruiting under-represented populations assisted by resources from the Penn Genome Frontiers Institute. Newly recruited faculty will be mentored in the proposed center by assisting the researchers in integrating into the interdisciplinary community and providing the junior faculty with pilot projects within the center's scientific programs. The center will serve as a key organization for coalescing systems biology activities at Penn and the greater Philadelphia area.

Table 2: Participating Team for Center for Synthetic Cell Programming

Senior Personnel	Expertise
Chen (BioEng)	Tissue engineering, cell biology, microfluidic engineering, mechanosensory stimulation
Eberwine (Pharm)	Neurobiology, molecular biology, RNA chemistry, cell culture
Gregory (Biology)	RNA biology, Nextgen sequencing
Janmey (Physiology)	Durotactic cell stimulation, neuroscience, bioengineering
Kim (Biology)	Computational Biology, comparative genomics, cell biology
Kumar (Mech Eng)	Robotics, hybrid dynamical systems, image analysis
Lukes (Mech Eng)	Fluid dynamics, macromolecule transport, molecular dynamics, simulations
Meaney (Bio Eng)	Molecular neuroengineering, neural regeneration and repair, stochastic modeling
Murray (Genetics)	Development, single cell genomics, cell imaging
Pappas (ElecEng)	Control systems, dynamical systems, biological systems identification
Raj (Bioeng)	Systems biology, development, single molecule imaging
Sul (Pharm)	Molecular imaging, biotechnology, neuroscience
Sundaram (Genetics)	Developmental biology, *C. elegans*, genetics
Wang (Pathology)	Computational Biology, algorithms, databases
Technical Staff	Expertise
Alex Chekholko	LINUX system administration, system programming
Adam Peritz	Nucleic acid chemistry

Research Project 10: Project Title and Purpose

Effects of Clinical Research Regulatory Auditing and Management on Compliance Findings. The goal is to see if there are year-to-year changes/improvements in the types and severity of any regulatory compliance gaps noted on research regulatory audit.

Duration of Project

1/1/2006 - 12/31/2009

Project Overview

Background
Over the past 5 years, the University of Pennsylvania has implemented a number of quality improvement initiatives aimed at enhancing research quality, oversight, and safety. A number of quality assurance studies of the effects of certain of these initiatives are underway. This currently proposed complementary project would greatly enhance other quality improvement initiatives by focusing not on individual efforts, rather on the larger picture of research regulatory compliance across a broad variety of clinical research studies. By using a research auditing program, coupled with compliance management plans, year-to-year assessments of overall compliance can be well characterized and compared to track progress made across the research enterprise, rather than narrowly focused in any one specific type of compliance support initiative such as training, standardized templates, SOPs, etc.

Methods
An audit plan, that clearly specifies how studies are selected for auditing, what will be audited, and how auditing findings will be captured and categorized, will be applied to clinical research studies at the University. Additionally, based on a severity scale applied to monitoring findings for a study, any applicable compliance management plan will be developed and applied. The auditing effort is supported by written procedures, data collection instruments, and an electronic database which stores and categorizes findings and tracks monitoring visits. Monitoring findings will be classified a minor, moderate, major, and serious (each with underlying definitions) and summarized by number of each type. Data will be displayed as absolute numbers of the type of compliance categorizations associated with each audited investigator. This data will be compared annually to assess any changes to the number and type of audit findings and the number of required management plans.

Principal Investigator

Gregg J. Fromell, MD
University of Pennsylvania, School of Medicine
150 Anatomy/Chemistry Bldg.
3620 Hamilton Walk
Philadelphia, PA 19104-6061

Other Participating Researchers

Nancy Pultorak - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

The expected outcome of this study will be that research teams will benefit from immediate feedback of compliance findings as well as collaborative management plan development. The institution will benefit by gaining clear knowledge of the type and level of research regulatory compliance, as well as a method to focus regulatory compliance management support for any findings. Research subjects will benefit by enhanced research oversight and management.

Summary of Research Completed

During the 2009 calendar year, the average number of active IRB- approved studies in the School of Medicine was 3185 (range: 3120-3294). Of those studies, an average of 1041 studies (range 1029-1060) received convened-board review (as opposed to an administrative review), an indication that the study involved greater than minimal risk activity. The routine audit program focused on the convened-board review studies. All studies undergoing a 6-month IRB review cycle (considered high risk) and studies for which the University faculty investigator held a FDA Investigational New Drug Application or Investigational Device Exemption responsibilities regardless of IRB-risk category, were audited. For the remainder of the research not falling into the above categories, the research audit focused on auditing a sample of studies for all faculty involved in convened-board, IRB-approved clinical research. For these investigators, the goal was to audit at least one full-board IRB-approved study per investigator, assuming the investigator's study was still actively enrolling and had enrolled at least one subject since the prior annual audit visit.

A total of 171 Investigators and 243 studies were audited. Audit findings were collected and assessed, and used to assign an overall compliance categorization to each audited study. There were no studies categorized as having serious non-compliance. Studies classified as having major and moderate non-compliance had findings that were procedural in nature and did not result in harm to subjects. A compliance management plan was developed for investigators with serious, major and moderate findings. Follow up auditing of all studies with major or moderate audit findings was conducted according to the study-specific management plan until successful completion of the requirements of the associated management plan. The following table summarizes the number investigators subgrouped by audit severity (an investigator was grouped by the most severe compliance category of all studies audited for that investigator).

Table 1. Summary of Compliance Findings – Routine Audit Program

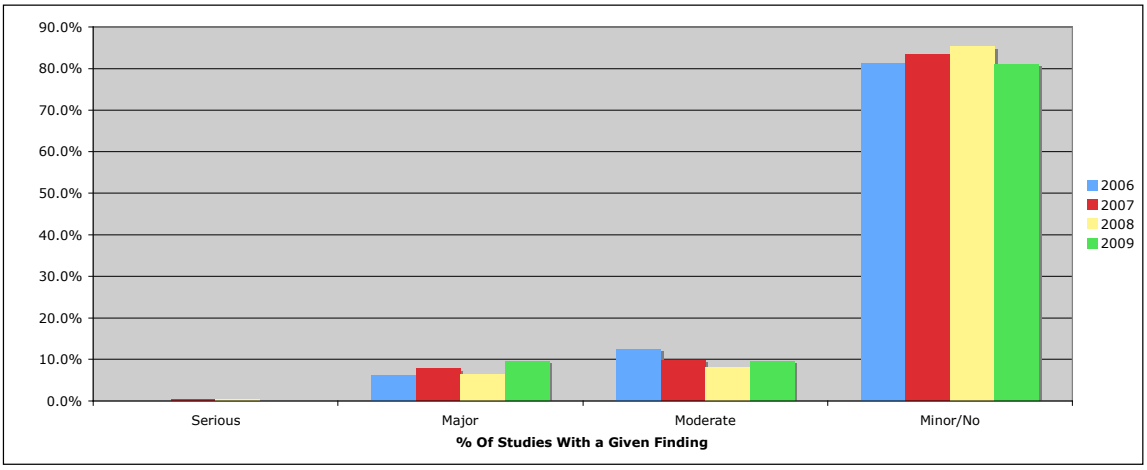
		Categories of Non-Compliance Findings* Counts of Investigators with Findings			
Total Investigators	Total Studies	Serious	Major	Moderate	Minor or No Findings
171	243	0	24 (14%)	11(%)	75 (%)

* Categories are mutually exclusive: investigators are included in a category based on the most significant of all findings for that investigator

The following graph displays the number of studies with specific audit categorizations as a percentage of all studies audited. Comparison of year-to-year results shows a continued high rate of studies with minor or no audit findings, though with a decrease in this category, and a corresponding increase in the categories of moderate and major non-compliance. There was a nominal increase in the number of studies with moderate findings, and a more prominent increase in the number of studies with major findings when compared to previous years. Two factors appear to be the most commonly associated with the increase in the incidence of major non-compliance: an increased number of investigators audited who had not been previously audited, and inadequate reporting of cases with exceptions to study inclusion/exclusion criteria.

In addition to reporting such findings to the IRB, non-compliance was addressed by a compliance management plan overseen by the Office of Human Research. These findings also allowed for relatively straightforward system solutions that included: direct feedback to investigators and their teams, feedback at the regular debriefing meetings with departmental chairs, provision of simple tools to further support documentation of inclusion and exclusion criteria, adjustment of training programs for investigators and coordinators, and points of emphasis in research e-newsletters. The results of the above interventions will be measured in the ensuing audit year. Concerning last year's reported results where findings of major non-compliance were mainly due to administrative errors in managing the HIPAA authorization separate from the consent form, and tardiness in filing annual reports for faculty with IND responsibilities, systemic interventions resulted in essentially complete resolution of those categories of non-compliance for this most recent audit cycle.

The auditing program has been a valuable endeavor that provides oversight of the conduct of clinical research as well a means to increase awareness of and attention to the wide range of investigator responsibilities. Insights from the auditing program have provided a means to inform and focus our research teams on areas for improvement, and also our training programs targeting clinical research investigators and coordinators and to develop tools to assist researchers with study management and associated regulatory compliance.



Research Project 11: Project Title and Purpose

Roles of ATR, H2AX and DNA-PKcs in Genome Maintenance and Tumor Suppression - Evidence suggests that mutations of the proteins ATR, H2AX and, likely, DNA-PKcs are risk factors for development of human cancers. Understanding the functional interaction between ATR and DNA-PK/H2AX (a signaling pathway) and their role in suppressing chromosomal translocations and the development of tumors has broad significance in cancer etiology, prognosis and treatment. This project examines whether two distinct pathways in genome maintenance may be linked, and whether or not these distinct pathways act cooperatively in suppressing the mutations and genomic rearrangements that cause cancer.

Duration of Project

6/1/2007 – 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 12: Project Title and Purpose

A Propensity Score Approach for Analyzing Data from Genetic Association Studies - The primary purpose of this project is to develop statistical methods that will allow cancer researchers to better estimate the effect of genes on the risk of cancer incidence. Statistical models that investigate the association between single nucleotide polymorphisms (SNPs) or multiple adjacent SNPs (haplotypes) with cancer incidence will be developed. These methods will allow the comparison of cases and controls and the presence and absence of a particular genotype or haplotype after taking into account the differences in non-genetic patient characteristics that may influence their cancer risk.

Duration of Project

1/1/2006 - 11/1/2009

Project Overview

In case-control genetic epidemiologic studies, regression methods can be used to estimate the association between a series of adjacent genetic polymorphisms and disease. In population-based genetic association studies, it is important to control for the impact of known or suspected patient and environmental factors because large differences in study subject characteristics, if not properly controlled for, can severely bias estimated genetic effects. One type of confounding bias that is unique to genetic studies is due to population stratification (PS) which arises when differences in both allele and disease frequencies exist in a population of mixed racial/ethnic subpopulations. Genomic control, structured association, principal components analysis (PCA),

and multidimensional scaling (MDS) approaches have been proposed to address this bias using genetic markers. However, confounding due to PS can also be due to non-genetic factors. Propensity scores are widely used to address confounding in observational studies but have not been adapted to deal with PS in genetic association studies. We propose a genomic propensity score (GPS) approach to correct for bias due to PS that considers both genetic and non-genetic factors. We compared the GPS method with PCA and MDS using simulation studies and found that GPS and MDS are comparable in terms of statistical properties such as bias, type I error, and power but with less computational burden. The GPS method provides a novel and robust tool for obtaining less-biased estimates of genetic associations that can consider both genetic and non-genetic factors.

Principal Investigator

Nandita Mitra, PhD
University of Pennsylvania
929 Blockley Hall
Department of Biostatistics and Epidemiology
423 Guardian Drive
Philadelphia, PA 19104

Other Participating Researchers

Timothy Rebbeck, PhD Huaqing Zhao, MS - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

The expected outcome of this research project is a set of statistical tools that will allow cancer genetics researchers to estimate the increase in risk of any particular type of cancer that can be attributed to a particular set of alleles in a gene. The developed set of methods will provide a novel and important analytical strategy to obtain less biased and more valid estimates of the impact of genes on cancer incidence by allowing researchers to incorporate important patient characteristics in their risk estimates. These methods will offer distinct advantages because they will allow one to control for numerous patients characteristics simultaneously which greatly simplifies statistical modeling. Moreover, this project will specifically benefit prostate cancer research by providing clinicians and geneticists a better understanding of the risk of CYP3A genes on prostate cancer incidence and severity. This information will help in understanding the underlying etiology and genetic basis of prostate cancer and will advance the study of this disease.

Summary of Research Completed

Background
There are a number of statistical methods that have been developed to test for and/or adjust for population stratification (PS) in population based genetic association studies. However, no consensus has been reached as to which method is best. These methods all utilize genotype information either from a set of random (null) markers or from a set of selected ancestry

informative markers (AIMs). We have developed a propensity score framework to optimally estimate and correct for the effect of genetic and non-genetic covariates in population-based association studies, thereby reducing the bias introduced by PS [Zhao H, Rebbeck T, Mitra N: A Propensity Score Approach to Correction of Bias due to Population Stratification Using Genetic and Non-Genetic Factors. *Genetic Epidemiology* 33(8): 679-690, 2009].

In our work we conducted simulations to investigate the statistical characteristics of our genetic propensity score (GPS) approach under several models. Specifically, we estimated bias, standard errors and coverage probabilities of GPS adjusted and unadjusted models. We also conducted extensive simulations to compare commonly used methods such as principal components (PCA) and multi-dimensional scaling (MDS) with our GPS approach. Our simulations demonstrated that GPS consistently performed as well or better than MDS or PCA. We now extend our methods to allow for three category genotypes (rather than the simpler case of a binary "treatment" or risk group) using an extended genetic propensity score (*eGPS*). This methodology estimates average genetic effects across three-level genotypes while still maintaining the advantages of the original propensity score.

In this report, we describe the extension of the propensity score. We then study the operating characteristics of our extended propensity score approach using simulations to estimate the effect of a genotype under an additive genetic model simultaneously adjusting for confounders such as random null markers, admixture informative markers (AIMs) and patient characteristics. We illustrate this approach in a case-control association study of testicular germ cell tumors and *KITLG* and *SPRY4* susceptibility genes.

Methods

We use G to denote the candidate genotype of interest, D to denote disease status, X to denote a vector of genetic covariates such as null markers and AIMs and non-genetic covariates which may include patient demographics and disease specific variables such as tumor characteristics. For all terms, the indices *i, k,* and *g* represent individual, subpopulation, and test-locus genotype, respectively. For example, p_{gik} represents the probability of having a high-risk genotype *g* for individual *i* in subpopulation *k*.

We define the *extended genomic propensity score* (*eGPS*) to be the likelihood of an individual having a particular test-locus genotype based on that individual's covariate makeup. This can be stated explicitly as

$$eGPS_i(\boldsymbol{g}_i\ ,\ \boldsymbol{x}_i) = \mathrm{P}(G_i|\ \boldsymbol{x}_i)$$

where $eGPS_i(\boldsymbol{g}_i\ ,\ \boldsymbol{x}_i)$ is the genomic propensity score for subject *i* calculated from that subject's $\boldsymbol{x}_i$ which represent that individual's vector of genetic and non-genetic covariates and where G_i is that subject's test-locus genotype. For example, for an additive disease risk allele A and reference allele a, G = 2 may represent the high-risk genotype *AA,* G = 1 may represent the heterozygote risk genotype *Aa,* and G = 0 may represent the homozygote wild-type genotype *aa*.

There are two important properties of the eGPS that play a vital role in the usefulness of the score in obtaining unbiased estimates of genetic associations. First, if group assignment is weakly ignorable given background covariates, then it is weakly ignorable given the extended

propensity score. Weak ignorability implies pairwise independence of the group with each of the potential outcomes, conditional on the observed covariates. Secondly, given the $eGPS(g, x)$, X and G are conditionally independent, i.e.,
$\Pr(X, G \mid eGPS(g,x)) = \Pr(X \mid eGPS(g,x)) \Pr(G \mid eGPS(g,x))$, which balances measured covariates across test-locus genotype groups (balancing property).

We define a general class of models that specify the potential relation among disease, test-locus genotype, genetic and non-genetic covariates.

$$f(E(D \mid X)) = \eta$$

where f (.) is a link function, such as the logit function, that determines the relationship between the outcome variable D and predictor variable X. E (D | X) denotes the conditional mean of D given X; and η is a function of covariates, usually a linear function. Finally, we specify our proposed eGPS model:

$$\eta = \alpha + \beta G + \gamma_1 eGPS(G=1, X) + \gamma_2 eGPS(G=2, X)$$

where α and β are log odds-ratios and γ1 and γ2 are nuisance parameters.

We fit the disease prevalence model assuming a logit link to estimate the effect of risk genotype(s) β :

$$\text{Logit}(D|G, eGPS(G, X)) = \alpha + \beta G + \gamma_1 eGPS(G=1, X) + \gamma_2 eGPS(G=2, X)$$

Simulation Results

We conducted simulation studies to evaluate the performance of the proposed eGPS method and to compare it with a model without adjustment for non-genetic and genetic covariates using polytomous logistic regression. First, to estimate the extended genomic propensity eGPS(X), we used polynomial logistic regression to obtain the predicted probability of having the number of risk alleles based on X. The dependent variable was the test-locus genotype rather than the disease status, and the independent variables are the covariates. Next, to estimate the effect of the genotype on disease, we used a second polynomial logistic regression model. In this case, the dependent variable was disease status, and the independent variables are the test-locus genotype variable along with the estimated eGPS.

For each scenario, we generated 1,000 simulation replicates and evaluated odds-ratio (OR), bias, standard error (SE) of the estimated log OR, the empirical coverage probability (CP), type I error and power.

Figures 1 and 2 illustrate that the eGPS retains close to nominal type I error rates (0.05) under most scenarios and has high power under all scenarios (>80%). Table I shows the results comparing no adjustment and eGPS methods assuming the true OR=1. Under no/mild PS, both methods performed well in terms of estimated OR and SE. However, our proposed eGPS method consistently outperforms the no adjustment method in terms of bias and CP. Under moderate PS, no adjustment yielded overestimation (OR ranging from 1.483 to 1.6) for all combinations of Fst, p1 and p2. Without adjustment, the coverage probability was almost zero with ranging from 0 to 0.003. However, our eGPS method performed consistently well in terms of estimated OR (ranging from .982 to 1.05), bias (close to 0) and CP(>95%). Similarly, under severe PS, no adjustment performed poorly, but the eGPS method consistently outperformed the no adjustment method in terms of estimation, bias and coverage probability.

Application

We used previously published data from a case-control genome-wide association study (GWAS) of testicular germ cell tumors (TGCT) to assess the utility of our eGPS approach [Kanetsky PA, Mitra N, Vardhanabhuti S et al: Common variation in KITLG and at 5q31.3 proximate to SPRY4 predispose to testicular germ cell cancer. *Nature Genetics* 41(7): 811-5, 2009]. To identify genes associated with TGCT development, the study performed a GWAS using cases comprised of 277 white, non-Hispanic men with pathologically-defined TGCT seen at the University of Pennsylvania Health System (UPHS) or Fox Chase Cancer Center (FCCC) in Philadelphia, PA. The study genotyped DNA extracted from venous blood using the Affymetrix® Genome-Wide Human SNP Array 6.0. The frequency of observed genotypes among TGCT cases were compared to those available from 919 white, non-Hispanic males from the Philadelphia region genotyped on the same Affymetrix platform. The most significant SNPs (those that reached statistical significance at a genome-wide threshold of $P < 5.0 \times 10^{-8}$ level) occurred within the *KITLG* (c-KIT ligand) gene region on 12q22, *SPRY4* (sprouty homolog 4) coding region on 5q31.3, and *SPRED2* (sprouty-related, EVH1 domain containing 2) on 2p14. The 1196 study samples used in the GWAS analyses were identified to fall within one multi-dimensional scaling (MDS) cluster indicating that these cases and controls were homogeneously of European ancestry. We analyzed these MDS identified 277 cases and 919 controls by further adjusting for population stratification using our eGPS method which utilized approximately 500 randomly selected null markers (~ 20 SNPs per chromosome) across the genome. Table 2 shows results from the original analysis and from the eGPS adjusted analysis. Note that the eGPS results attenuate the findings slightly towards the null indicating that some bias from residual population stratification may have remained after using MDS and that eGPS served to produce less biased effect estimates.

Conclusions

We conclude that our method provides a novel and broadly applicable analytic strategy for obtaining less biased and more valid estimates of genetic associations under a broad range of genetic models.

Table 1. **Estimated Odds-ratio (OR), Bias and Standard Error (SE) of the Estimated log OR and the Coverage Probability (CP) for Comparing no Adjustment and eGPS Methods with True OR=1**

			Estimate OR		Bias		SE		CP	
Fst	p1	p2	No Adjustment	eGPS	No Adjustment	eGPS	No Adjustment	eGPS	No Adjustment	eGPS
No/ Mild Population Stratification										
0.01	0.1	0.3	1.207	0.990	0.185	-0.015	0.089	0.101	0.456	0.964
0.01	0.1	0.5	1.137	1.024	0.127	0.018	0.079	0.111	0.685	0.967
0.05	0.1	0.3	1.180	0.958	0.162	-0.048	0.088	0.103	0.537	0.926
0.05	0.1	0.5	1.135	1.003	0.125	-0.002	0.079	0.114	0.701	0.964
0.1	0.1	0.3	1.127	1.007	0.117	0.002	0.081	0.107	0.746	0.963
0.1	0.1	0.5	1.116	0.994	0.108	-0.011	0.080	0.110	0.788	0.969
Moderate Population Stratification										
0.01	0.1	0.3	1.483	1.050	0.390	0.044	0.091	0.104	0.003	0.937
0.01	0.1	0.5	1.587	1.041	0.460	0.035	0.080	0.114	0.000	0.962
0.05	0.1	0.3	1.486	0.982	0.393	-0.024	0.089	0.106	0.004	0.953
0.05	0.1	0.5	1.600	1.021	0.468	0.015	0.080	0.117	0.000	0.958
0.1	0.1	0.3	1.538	1.015	0.428	0.009	0.082	0.110	0.000	0.974
0.1	0.1	0.5	1.547	1.012	0.434	0.006	0.082	0.113	0.000	0.958
Severe Population Stratification										
0.01	0.1	0.3	2.050	1.174	0.714	0.155	0.094	0.111	0.000	0.728
0.01	0.1	0.5	2.753	1.079	1.010	0.068	0.088	0.127	0.000	0.930
0.05	0.1	0.3	2.145	1.030	0.759	0.024	0.093	0.117	0.000	0.950
0.05	0.1	0.5	2.832	1.047	1.039	0.037	0.089	0.132	0.000	0.948
0.1	0.1	0.3	2.555	1.032	0.935	0.025	0.089	0.123	0.000	0.961
0.1	0.1	0.5	2.653	1.048	0.973	0.040	0.089	0.126	0.000	0.951

OR=odds ratio between disease (D) and gene (G); estimated OR, bias and SE are the average over 1,000 replications of the estimated OR, bias and SE of the estimated log OR; bias=log(estimated OR)-log(true OR); the empirical coverage probability is the proportion of the confidence interval of the estimated OR including the true OR among 1,000 replications. In all models, true OR=1.

Table 2: Risk Estimates for SPRED2, SPRY4 and KITLG SNPs for MDS identified Caucasian samples with and without further eGPS adjustment (277 cases and 919 controls)

Gene	Marker	Per allele OR (95% CI)	
	(non-risk/risk)	*No Adjustment*	*eGPS*
SPRED2	rs17031166 G/C	1.60 (1.31 – 1.96)	1.46 (1.14 – 1.88)
	rs1549383 A/G	1.65 (1.35 – 2.02)	1.57 (1.23 – 2.01)
SPRY4	rs4324715 T/C	1.59 (1.30 – 1.93)	1.48 (1.16 – 1.88)
	rs1549383 C/T	1.59 (1.31 – 1.94)	1.45 (1.14 – 1.85)
KITLG	rs3782179 A/G	2.36 (1.73 – 3.21)	1.97 (1.29 – 3.01)
	rs4474514 A/G	2.45 (1.79 – 3.35)	1.97 (1.29 – 3.01)

Figure 1 Type I error of eGPS by MAF and OR. Dotted line is no/mild PS. Short dashed line is moderate PS. Long dashed line is severe PS.

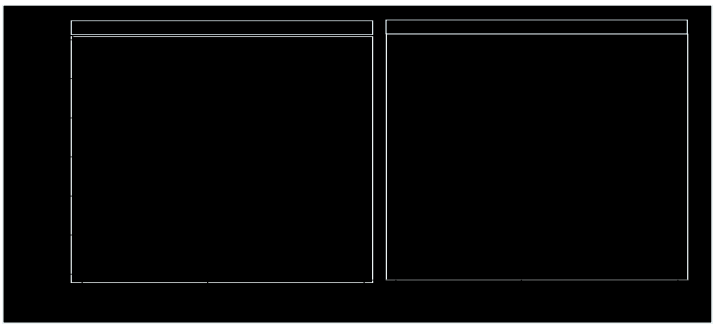
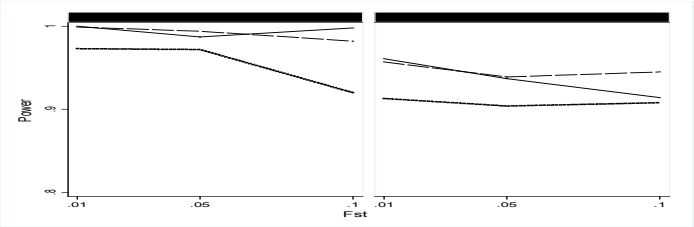


Figure 2 Power of eGPS by MAF and OR. Dotted line is no/mild PS. Short dashed line is moderate PS. Long dashed line is severe PS.



Research Project 13: Project Title and Purpose

Anti-Her2/neu Polymersomes for Breast Cancer Diagnosis and Therapy - Early detection of breast cancer is critical for successful treatment. Tumor detection by non-radioactive, optical imaging is a new opportunity to detect and diagnose early breast cancers in a cost-effective manner. Her2/neu and EGF receptors are validated therapeutic targets for breast cancer therapy. Anti-Her2/neu peptidomimetics (AHNP) are small antibody-like molecules that bind to Her2/neu receptor and limit tumor growth in mice similar to the anti-Her2/neu monoclonal antibody, Herceptin® (Genentech, Inc). In this project, nanoparticles consisting of the small anti-Her2/neu antibody-like peptidomimetic (AHNP) conjugated to polymersomes with near-infrared emitting fluorescent (NIRF) probes will be used to visualize breast tumors in mice. In addition, the therapeutic efficacy of the anti-Her2/neu nanoparticle will be studied in mice.

Duration of Project

1/1/2006 - 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 14: Project Title and Purpose

Role of Daxx in Tumorigenesis - The tumor suppressor p53 exerts its anti-proliferation function through the induction of growth arrest and apoptosis in response to various cellular stresses. Nearly 50% of human tumors have mutations in the p53 gene, and in the tumors with wild type p53, its function is often compromised through the alterations in its regulatory proteins or targets. The function of p53 is inhibited by Mdm2, and the inactivation of p53 can be accomplished by the overexpression of Mdm2 in tumors. We have recently found that a cellular protein known as

Daxx is required for the function of Mdm2 and that Daxx may also be overexpressed in a substantial number of tumors. This project will investigate the possibility that Daxx is an oncogene through the inhibition of p53. This work will likely reveal an important regulatory mechanism of p53 and may identify a novel target for cancer therapy.

Duration of Project

1/1/2006 - 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 15: Project Title and Purpose

Identification of Barriers and Facilitators to Oncology Clinical Trials Recruitment - It is well documented that only about 3-5% of new cancer patients participate in clinical trials and that almost 80% of adult studies suffer from recruitment issues. Despite considerable time, effort and expense, the proportion of patients recruited is frequently small. This study will assess barriers and facilitators to clinical trials recruitment at the University of Pennsylvania within the context of patient-, clinician- and organizational-level factors. The data will provide benchmarks against which to evaluate the effectiveness of future interventions and evidence upon which to develop interventions for improvement of recruitment to oncology clinical trials.

Duration of Project

6/12/2006 – 12/31/2009

Project Overview

Much observation, yet little scientific rigor, has been applied to the study of barriers to and methods for increasing recruitment to cancer clinical trials. However, without this obvious prerequisite to successful clinical trials completion even the most innovative research is stymied. The 2005 Agency for Healthcare Research and Quality Report on recruitment to clinical trials found only nine studies of physician barriers and facilitators to clinical trials recruitment. Most of the studies found physician attitudes to be barriers and only one study found physician attitudes to facilitate clinical trials recruitment.

In order to improve recruitment and retention to clinical trials, the University of Pennsylvania Abramson Cancer Center is committed to a developmental Recruitment, Retention and Outreach Core (RROC) Facility to study barriers and rigorously test interventions to facilitate informed decisions and recruitment.

The primary aim of this project is to assess clinician (physician and clinical research associates CRAs)-level barriers and facilitators to recruitment to cancer clinical trials at the University of Pennsylvania (Penn) Abramson Cancer Center. *Hypothesis: Based on pilot work conducted by the PI through the Radiation Therapy Oncology Group, we hypothesize that there will be significant differences between physician and CRA beliefs, attitudes and practices regarding clinical trials and clinical trials recruitment. This discord will help identify specific barriers to - clinical trials and further inform hypotheses for education and communication training interventions that may improve recruitment.* The secondary aims are 1) to assess organizational-level barriers and facilitators to recruitment to cancer clinical trials, 2) to assess patient-level barriers to clinical trials. Penn cancer clinical trials accrual by diagnosis, stage, race, gender, age and when available distance living from hospital, will be compared to United States Census and Surveillance Epidemiology and End Results (SEER) data as well as primary catchments statistics, and 3) to compare accrual by specific trial variables (Phase, cancer type, and whether it is a treatment, symptom management, behavioral or cancer prevention trial) and by funding source/coordinating entity (pharmaceutical company, National Institute of Health, cooperative group, investigator initiated, etc). Data gathered will be used as pilot data for future funding for studies of specific interventions to address the barriers and augment the facilitators identified in this study.

Principal Investigator

Deborah Watkins Bruner, RN, PhD
University of Pennsylvania
Nursing Education Building, Room 352
420 Guardian Drive
Philadelphia, PA 19104-6096

Other Participating Researchers

Carmen E. Guerra, MD, Lynn Schuchter, MD, Connie Ulrich, RN, PhD - employed by University of Pennsylvania

Expected Research Outcomes and Benefits

Clinical trials provide evidence that is formulated into recommendations for practice guidelines, but without improvements in recruitment it is an ongoing concern that patients enrolled in trials may not be similar to those treated in routine practice. There is some evidence that clinical trials have a positive effect on patient outcomes and therefore, for reasons of social justice, all cancer patients should have access to the high-quality care, surveillance and availability of the latest treatments put forth in clinical trials. Identifying barriers to recruitment into randomized clinical trials can help researchers adjust recruitment strategies to maximize enrollment. Data gathered from the assessments will be used as pilot data for future funding for studies of specific interventions to address the barriers and augment the facilitators of recruitment identified in this study. For example, if communication is found to be a major barrier to recruitment, scripted informed consent and or video-taped examples of effective clinical trials presentation and informed consent could be developed and tested. If patient socio-demographic factors are a

barrier, specific outreach and community partnerships may assist in recruitment. And if organizational barriers, such as CRA method for identification of eligible patients, are identified, improved methods and CRA training may be a solution that could be tested. This work is a priority for Penn's Abramson Cancer Center as demonstrated by the work of the Clinical Trials Working Group with its focus on improving minority accrual and the recent development of a Recruitment, Retention and Outreach Core (RROC) facility which will ultimately focus on prospectively designing recruitment plans for developing clinical trials.

Summary of Research Completed

Organizational-level barriers and facilitators – *Reported previously.*

Clinician-level barrier and facilitators -Focus groups
As of this time, coding and qualitative analysis for 5 focus groups and 3 individual interviews are in progress and a manuscript is in preparation for *Contemporary Clinical Trials.*

Clinician-level barrier and facilitators – Survey
We developed a survey to assess clinician level barriers and facilitators to clinical trials based on a previously pilot-tested survey in the national clinical trials Radiation Therapy Oncology Group (RTOG). The survey included items on knowledge, attitudes and communication styles about clinical trials/research. The survey was revised based on an updated review of literature and to accommodate sensitivity to detecting facilitators and barriers potentially unique to a hospital within an academic institution (e.g., a PhD investigator without clinic appointment may have variables different from MD investigators with primary appointment in a hospital/clinic). Drafts were reviewed by members of the Penn Clinical Trials Working Group.

The survey was uploaded on to SurveyMonkey.com, enabling a secure web version of the survey. The online survey was checked for its operative function by two research staffs and corrected for any logistical problems. The survey was also set to ignore the recording of recipient identifying information (e.g., names and email addresses) to suit the anonymous nature of the survey participation.

Investigators, physicians, and research staffs at the University of Pennsylvania who were involved with cancer-related clinical research with human subjects were identified as a target population. The list of potential respondents was constructed from three sources that cast a wide net over faculty and staffs involved in cancer clinical research across Penn: Penn Clinical Trials Scientific Review and Monitoring Committee (CTSRMC), Abramson Cancer Center's clinic and personnel web listings, and Abramson Cancer Center membership. Altogether, the list included not only physician-investigators and clinic staffs involved in cancer research but also physicians who are involved as referring physicians and PhD- and MD-level investigators and staffs external to hospital departments who are involved in cancer-related clinical research with either leading or collaborative capacity.

The initial total of main list contained 1089 names; 493 duplicates and invalid status holders (i.e. no longer employed at Penn) were removed, and 37 emails were returned as undeliverable at the

time of invitation to respond to an online survey. This process left a total of 559 individuals who were invited by email to take the online survey.

A total of 201 online surveys were returned, which included 100 completed entries. The overall response rate was 36%, and 18% for completed entries. Prospective respondents were e-mailed an initial email from the Director of the Cancer Center and two additional email reminders from the study PI to take the survey in two-week intervals. The survey was open for six weeks.

Of respondents who completed the survey, generally there were slightly more women (58%). 43% were physician investigators (MD PI), 20% were clinical research associates (CRA), 34% were non-physician investigators (non-MD PI), and 3% were referring physicians (MD). The gender of respondents varied with occupational role in cancer clinical research: 55% CRAs, 24% MD PIs, 19% non-MD PIs, and 2% MD for women and 5% CRAs, 69% MD-PIs, 21% non-MD PIs, and 5% MD for men. 79% of completing respondents were White, 7% were of African American and 15% were of Asian background. Average years in the involvement with cancer-related clinical research were 10.5 in general and 7.5 at the University of Pennsylvania.

While there was a total of 201 survey entries the amount of missing data per item made us determine that MD – CRA comparisons are not feasible. For items completed some useful data was collected that can help inform infrastructure and education. For example, while 70% respondents indicated they routinely screen patients for clinical trials, only 56% have a formal process or routine mechanism for screening. Best practices methods for screening patients for eligibility may be facilitated through the standing clinical research associates meetings. Almost 50% of respondents agreed that a regular accrual report would be an incentive for them to improve accrual and 17% were unsure. A system for regular feedback to investigators and clinical research associates may improve reaching accrual goals. Educational efforts need to include review of equipoise and clinical trials since only 46% of respondents endorsed the "uncertain" response to the question, "In general most patients will do as well with standard therapy as on a clinical trial." Although, 75% indicate they frequently or always discuss with patients the fact that there is uncertainty in which treatment is best. Benefits to clinical trials participation frequently or always discussed with patients include altruism (opportunity to help other patients in the future), endorsed by 92% of respondents and heightened surveillance, but this was only endorsed by 51% of respondents and may be improved with educational efforts.

The most common reasons perceived by investigators and research associates as why patients refuse to participation in clinical trials were fear of experimentation with 78% of respondents endorsing this at a level of 1-4 on a scale of 1-10 with 1 being the most common reason and 10 the least. Fear of receiving placebo and distance the patient has to travel to the hospital were each endorsed by 52% of respondents at the same level (1-4). Interestingly, although distance was highly endorsed as a barrier to accrual by 52%, only 30% of respondents thought lack of transportation was and a barrier. The perception that lack of health insurance is a barrier was not endorsed as a common barrier and, in fact, was perceived as the least common barrier with 78% of respondents ranking this issue as 5-9 on a scale of 1-10 with 1 being the most common reason and 10 the least.. Not having a provider of similar racial/ethnic background was not perceived as a major barrier by 70% of respondents and 91% thought that minority patients were very or somewhat willing to participate in clinical trials. Only 17% perceived language issues as a

barrier to clinical trials accrual although 35% thought their patients would benefit from translated consent forms.

Patient-level barriers and facilitators

Patient accrual by disease site is reported under organization-level barriers above. Our analysis of minority and gender clinical trials recruitment shows an under-representation of minorities by comparison to our catchments and to our health system in/outpatient population. In 2008 we accrued just under 15% minorities (Table 1) to therapeutic trials while our catchments that account for approximately 85% of the population surrounding our health system and also 85% of the population we service included 30% under-served minorities. In addition, while our 85% service area includes slightly more than half females, women account for only about 1/3 of our therapeutic clinical trials accrual. This information has provide rationale for a specific focus on minority recruitment within the newly developed Recruitment, Outcomes and Assessment Resource (ROAR). Among multiple goals, ROAR is designed to aid research by enhancing collaborative, multidisciplinary population research; enabling observational, behavioral, clinical translation and interventional studies; avoiding inefficiency in the development and execution of these studies; enhancing recruitment of minorities and other underrepresented groups.

In addition, it has been our intention to report on the same data for non-therapeutic (cancer control and population studies), however, we encountered barriers to assessing overall and minority accrual data to non-therapeutic trials since historically the data on these trials had not been tracked centrally. The cancer center has purchased a new clinical trials data management system and it now will include population and cancer control trials.

Table 1. Enrollment by gender and minority status by gender to therapeutic cancer clinical trials at the Abramson Cancer Center from 07/01/2007-06/30/2008.

GRAND TOTAL ENROLLED			749	
GRAND TOTAL MINORITY ENROLLED*			110	
GRAND TOTAL MALE MINORITY ENROLLED*			70	
OVERALL % MINORITY MALE				9.38
GRAND TOTAL FEMALE MINORITY ENROLLED*			40	
OVERALL % MINORITY FEMALE				5.37
OVERALL % MINORITY				**14.75**

Research Project 16: Project Title and Purpose

Lactic Acid as a Predictor for Severe Sepsis in the Hematologic Malignancies - To determine if serum lactic acid levels are predictive of severe sepsis in the hospitalized, hematologic malignancy patient population with a strong clinical suspicion of infection.

Duration of Project

6/12/2006 – 6/30/2008

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement Annual C.U.R.E. Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.