Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** Geisinger Clinic

2. **Reporting Period (start and end date of grant award period):** 1/1/2012-6/30/2013

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Victor Vogel, MD

4. **Grant Contact Person's Telephone Number:** (570) 214-9391

5. **Grant SAP Number:** 4100057661

6. **Project Number and Title of Research Project:** 1 -Reducing the Burden of Breast Biopsy with Abnormal Screening Mammograms

7. **Start and End Date of Research Project:** 1/1/2012-6/30/2013

8. **Name of Principal Investigator for the Research Project:** Victor Vogel, MD

9. **Research Project Expenses**.

   9(A) Please provide the total amount of health research grant funds spent on this project for the entire duration of the grant, including indirect costs and any interest earned that was spent:

   $ $61,035.54

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name, First Name	Position Title	% of Effort on Project	Cost (Salary & Benefits)
Vogel, Victor MD	Principal Investigator	2%	$5,888.43
Stark, Azadeh, PhD	Co-Investigator	15%	$29,912.07
Yan, Xiaowei	Co-Investigator	1.2%	$1,530.05

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name, First Name	Position Title	% of Effort on Project
None		

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

Type of Scientific Equipment	Value Derived	Cost
None		

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes_____ No____X_____

If yes, please indicate the source and amount of other funds:

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes__ No_____X_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

A. Title of research project on grant application	B. Funding agency (check those that apply)	C. Month and Year Submitted	D. Amount of funds requested:	E. Amount of funds to be awarded:
None	☐NIH ☐ Other federal (specify:_____ _____) ☐ Nonfederal source (specify: _____)		$	$

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes_____ No_____X_____

If yes, please describe your plans:

12. Future of Research Project. What are the future plans for this research project?

Currently, we do not have any research plan for the future.

13. New Investigator Training and Development. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes_____ No____X_____

If yes, how many students? Please specify in the tables below:

	Undergraduate	Masters	Pre-doc	Post-doc
Male				
Female				
Unknown				
Total				

	Undergraduate	Masters	Pre-doc	Post-doc
Hispanic				
Non-Hispanic				
Unknown				
Total				

	Undergraduate	Masters	Pre-doc	Post-doc
White				
Black				
Asian				
Other				
Unknown				
Total				

14. Recruitment of Out-of–State Researchers. Did you bring researchers into Pennsylvania to carry out this research project?

Yes_____ No_____X____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

15. Impact on Research Capacity and Quality. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes_____ No_____X___

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research.

16. Collaboration, business and community involvement.

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes_____ No___X__

If yes, please describe the collaborations:

16(B) Did the research project result in commercial development of any research products?

Yes_____ No_____X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No_____X____

If yes, please describe involvement with community groups that resulted from the research project:

17. Progress in Achieving Research Goals, Objectives and Aims.

List the project goals, objectives and specific aims (as contained in the grant agreement). Summarize the progress made in achieving these goals, objectives and aims for the period that the project was funded (i.e., from project start date through end date). Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.

Summary of Research Completed

We conducted a project with the primary objective of obtaining descriptive information about premenopausal women, defined as women younger than age 50, utilizing screening mammography at Geisinger Health System (GSH). As of 2009, GHS provided health services to a total of 1,188,388 women of whom 300,068 (25.25%) were between ages of 18 and 50 years. We differentiated screening from diagnostic mammography by restricting our search to mammograms limited to two images (craniocaudal and mediolateral oblique of each breast) and/or if the radiologist has indicated the procedure as such. A random group of women (N=5,100) between ages of 18 and 49 who had utilized screening mammography between January of 2010 and December of 2011 were selected. Details of the population's age, family history of breast cancer (BC), presence of BRCA 1/2 mutations and history of breast biopsy are shown in Table 1. Of the 218 women who underwent breast biopsy 175 (80.3%) were diagnosed with benign conditions of the breast and 5 (2.3%) with proliferative with atypia (2 focal with mild atypia). One woman was diagnosed with B cell lymphoma and one with metastatic melanoma to the breast. The remaining 26 (11.9%) were diagnosed with neoplastic lesion of the breast: 14 women were diagnosed with in situ lesions (10 ductal & 4 lobular) while12 women were diagnosed with invasive BC.

Table 1: Description of the sampled screening mammography utilizers ≤ 49 years of age

Descriptive Variable	N (%)
Age	
18-29	63 (1.2)
30-39	361 (7.1)
40-49	4,767 (91.7
Family History of BC	
Yes	1,895 (37.2)
No	3,205(62.8)
BRCA1 &2 Tested	
Yes	24 (0.47)
No	5,076 (99.53)
MRI	
Yes	19(0.37)
No	5,081 (99.63)
Biopsy	
Yes	218 (4.3)
No	4.882 (95.7)

MyCodeBiobanking: The Geisinger MyCode population-based cohort project is a prospective repository of blood samples (DNA and serum) that was implemented in early 2007. MyCode project is a major resource for research that combines information obtained from DNA and serum with health information from the e-health records and other e-resources intended to improve the prevention, diagnosis, and treatment of disease. The well annotated banked samples are representative of the 43 county service catchments of GHS because of the high acceptance rate (89%). Currently, a total of 23,000 individuals of whom 2,788 are women ≤ 49 years of age have donated blood samples to the MyCode biobanking project. At the time of collection, blood samples are processed according to the protocol, serum and DNA are then aliquoted into freezer vials and identified by unique identifiers before banking in the designated freezers. All samples can be linked to various e-databases should an investigator require data. The MyCode cohort project is in full compliance with the U.S. Congress Health Insurance Portability and Accountability Act (HIPAA) of 1996 and has the approval of the Institutional Review Board.

Data Element
The primary source of clinical data is Geisinger's Clinical Decision Intelligence System (CDIS), an enterprise-wide warehouse where data are stored at the most granular level to allow for unlimited reporting and analysis. The CDIS consists of:1) Geisinger Electronic Health Records, including digital information that dates back to 2004; 2) Insurance claims; 3) Billing Information; 4) Patient Safety; 5) Surgical Pathology; 6) Laboratory Medicine; 7) Pharmacy (in & out patients); 8) Radiology and 9) Patient Satisfaction. To comply with internal policy and HIPAA the Division of Biostatistics Research and Data Core (BRDC) has been designated as a data broker. The data broker designation allows direct access to clinical data by the BRDC using standard statistical software and will greatly enhance and expedite the research process timeline. A key role of the broker function is the linkage of biological specimens to longitudinal clinical data.

Selection of SNPs
In selecting the genes and their SNPs we reviewed findings from GWAS and other independent studies and applied minor allele frequency filtering approach and function prediction method to select a total of 64 SNPs of eight genes (Frayling et al. 2007; Kakamani et al. 2011; Hunter et al. 2007; Dossus et al. 2008; Langsenlehner et al. 2006; Zhang et al.; Healey et al. 2011; Dossus et al. 2010; Stacey et al. 2007; Andreasen et al. 2008; Rebbeck et al. 2009; Easton et al. 2007; Brasky et al. 2011) (Table 2).

Table 2 List of the genes and their variants (SNPs) evaluated

Gene	Chromosomal Location	SNPs
FGGR2	10q26	rs 1219648, rs 11200014, rs 2981579
COX-2/PTGS2	1q25.2-25.3	rs 2745559, rs 689470, rs 689466, rs 2206593, rs 5277, rs 12042763, rs 2383529
FTO	16q12.2	rs 9939609, rs 1861868, rs 1477196
GHRL	3q26.3	rs 171336, rs 171407, rs 35684, rs 4684677, rs 2075356, rs 696217, rs 27647, rs 3755777, rs 27498, rs10490815
GHSR	3q26.2	rs 2948694, rs 2922126
IL6	7p21.0	rs 4552807 rs 6969502 rs 6952003 rs 10156056 rs 7776857 rs 7801617 rs 7805828 rs 12700386 rs 1800795 rs 2069840 rs 2069861 rs 10242595 rs 11766273
MAP3K1	5q11.2	rs 889312
ESRα	6q25.1	rs 2046210, rs 12662670, rs 3020314

Laboratory analysis
Banked samples were retrieved and were sent to the core laboratory for analysis. All samples

were marked with the study unique identifiers and the laboratory personnel and the collaborating investigators remained masked to the status of samples.

DNA isolation

DNA was extracted from EDTA-anticoagulated whole blood using QIAsymphony SP Robot with Qiagen QIAsymphony DNA Midi Kit (Qiagen, Valencia, California) according to the manufacturer's protocol. Quantification of extracted DNA was performed using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, Delaware).

Genotype analysis

Single nucleotide polymorphism genotyping was performed on TaqMan® OpenArray System with assay kit (64 assay format) and Genotyping Master Mix purchased from Life Technologies (Life Technologies, Foster City, California), according to the manufacturer's protocol. Briefly, 10 ul of each DNA samples (containing 10 ng of DNA, 5 μL of TaqMan Genotyping Master Mix, 0.25 μL of 40x assay mix, and water) plated in 384 well plate were loaded on OpenArray assay slide with Life Technologies OpenArray® AccuFill™ System (Life Technologies, Foster City, California) then performed PCR on GeneAmp PCR System 9700 (Life Technologies, Foster City, California) as follows: 93 °C for 10 minutes followed by 50 cycles at 95 °C for 45 seconds, 94 °C for 13 seconds, and 53 °C for 2 minutes 14 seconds. The post-PCR OppenArray assay slides were then scanned with OpenArray scanner and analyzed using TaqMan genotyper Software v1.3 (Life Technologies, Foster City, California). We took a two-step quality control measure to remove poor quality genotype data. First, 10 % samples were replicated to test the concordance and reliability of the genotyping result. We excluded discordant SNPs. This step was followed by excluding SNPs with a recall rate of < 85 % for genotyping; this step was followed by manual recall for the remaining SNPs. A total of 40 SNPs passed the two-step quality control requirement. .

Linkage disequilibrium and haplotype analysis

The observed frequencies for all selected SNPs in our sample were compared with and were in agreement with the Hardy-Weinberg-Equilibrium. We then evaluated the linkage disequilibrium structure of the SNPs in our sample using the Gabriel algorithm (Gabriel et al. 2002). (HaploView 4.0 Day Lab, Cambridge, MA). This step is followed by reconstruction of the haplotypes to evaluate the interaction between SNPs. We conducted haplotype analysis using haplo-stats Version 1.4.0 (Sinnwell, JP and Schaid DJ, built in R, version 2.7.1). In this package the maximum likelihood estimate of a haplotype probability is calculated using the EM algorithm, and used to determine possible haplotypes.

Statistical analysis

Distributions of demographic and clinico-pathology variables between cases and controls were evaluated using non-parametric and parametric statistics. In developing the multivariate logistic regression model to determine the variables that were associated with the risk of PMBC, we first estimated the individual effect of each variable and their interactions with the outcome of interest, breast cancer. Variables with a P-value ≤ 0.10 were considered as the candidate variables. Interactions between variables also were tested at P-value $\leq .05$. The final model included five candidate variables (age, smoking status, alcohol consumption status, family history of breast cancer, and use of hormone replacement therapy (HRT). In our next analysis,

we restricted the reference group to controls with no history of exposure to HRT or smoking and no clinical documentation of MetS The final model included age, family history of breast cancer, HRT, MetS and the interaction between HRT and MetS. The estimated risk of PMBC was not significantly different from our first approach where all controls were inclusive. Therefore, we use this reference group to estimate the relative risk contributions of genetic polymorphism to PMBC in presence of clinico-demographic risk factors. For each SNP, testing each SNP individually for its association with PMBC, we used the Cockerham genetic model additive coding scheme and dominant coding scheme (Cordell 2002). For the additive coding approach, we assigned the zero, one or two to each SNP genotype according to the number of copies of minor alleles. For the dominant coding scheme, we assigned the value of one for rare homozygozity and zero for the alternative homozygotes. The SNPs which showed significant association by either coding scheme, were selected (P-value < 0.1). The final multivariable model was restricted to the dominant coding scheme and was adjusted by age, family history of breast cancer, HRT, MetS and the interaction between HRT and MetS. Finally, we evaluated the risk prediction ability of the final model by plotting the receiver operating characteristic (ROC) curves and calculated area under the curve (AUC), which was equivalent to c-statistics, and reported for each model.

We have completed statistical analyses and prepared final data interpretation. We have applied findings from this pilot project toward drafting a manuscript.

18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
____x_No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
___x__No

If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

___Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

____Number of subjects originally targeted to be included in the study
____Number of subjects enrolled in the study

Note: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females
_____Unknown

Ethnicity:
_____Latinos or Hispanics
_____Not Latinos or Hispanics
_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

19. Human Embryonic Stem Cell Research. Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
__X__No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?

_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

20. Articles Submitted to Peer-Reviewed Publications.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, and an abbreviated title of the publication. For example, if you submit two publications for Smith (PI for Project 01), one publication for Zhang (PI for Project 03), and one publication for Bates (PI for Project 04), the filenames would be:

Project 01 – Smith – Three cases of isolated
Project 01 – Smith – Investigation of NEB1 deletions
Project 03 – Zhang – Molecular profiling of aromatase
Project 04 – Bates – Neonatal intensive care

If the publication is not available electronically, provide 5 paper copies of the publication.

Note: The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

Title of Journal Article:	Authors:	Name of Peer-reviewed Publication:	Month and Year Submitted:	Publication Status (check appropriate box below):
1. None				☐Submitted ☐Accepted ☐Published

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes_____ No_____x____

If yes, please describe your plans:

21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project. Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

Findings from this project had the potential of impacting screening for lung and breast cancer in women; however, due to the poor quality of banked serum samples, we were not able to collect any information on the serum autoantibodies to breast and lung cancer.

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None.

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_____ No___X____

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No___

If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No___
If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No ___

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes___ No____

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes_____ No_X

If yes, please describe your plans:

24. Key Investigator Qualifications. Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

Victor G. Vogel, MD, MHS, Director of Geisinger's Cancer Service Line, he has more than 20 years of experience leading large focused projects, specifically clinical research in the use of selective estrogen receptor modulators for breast cancer risk reduction in women at increased risk. Dr. Vogel served as protocol chair for the National Surgical Adjuvant Breast and Bowel Project's Study of Tamoxifen and Raloxifene, a study that enrolled more than 19,000 postmenopausal subjects at increased risk of breast cancer and followed them prospectively for 7 years at more than 400 clinical sites in the US and Canada.