Award Number: W81XWH-05-1-0314

TITLE: Early life processes, endocrine mediators and number of susceptible cells in relation to breast cancer risk.

PRINCIPAL INVESTIGATOR: Dimitrios Trichopoulos, MD, Ph.D.

CONTRACTING ORGANIZATION: Harvard Medical School
Boston MA 02115-6027

REPORT DATE: April 2007

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
**Title and Subtitle:**
Early life processes, endocrine mediators and number of susceptible cells in relation to breast cancer risk.

**Authors:**
Dimitrios Trichopoulos, MD, Ph.D.

**E-Mail:** dtrichop@hsph.harvard.edu

**Abstract:**
Scope: To investigate the role of early life processes, endocrine mediators and number of susceptible cells on adult life breast cancer risk. Method: Five interlinked component projects covering the spectrum from endometrial to adult life. Progress report: Component projects 1 to 4 were officially launched July 2005. Component projects 5a and 5b were officially launched July 18, 2006. Tasks and subtasks to be performed were described in the submitted Statement of Work (SOW). Subtasks 4a, 4b, 5a, 5b have been completed. Subtasks 1a, 1b, 2a, 2b, 3a, 3b, 3c, 5c, 5d are ongoing. Subtasks 2d and 4c have been initiated. Subtasks 6a and 6b are being implemented. Major findings: Data collection is still on-going. Analyses are pending and no findings can be reported yet.

**Subject Terms:**
Breast cancer, early life, mammary gland specific stem cells, hormones
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Body</td>
<td>4</td>
</tr>
<tr>
<td>Key Research Accomplishments</td>
<td>13</td>
</tr>
<tr>
<td>Reportable Outcomes</td>
<td>14</td>
</tr>
<tr>
<td>Conclusions</td>
<td>14</td>
</tr>
<tr>
<td>References</td>
<td>14</td>
</tr>
<tr>
<td>Appendices</td>
<td>14</td>
</tr>
</tbody>
</table>
INTRODUCTION

The aim of the project is to investigate the role of early life processes, endocrine mediators and number of susceptible cells on adult life breast cancer risk. Based on the hypothesis that breast cancer risk is a function of number of mammary gland cells at risk of transformation and that this number is largely modulated by perinatal events and conditions, five component projects have been initiated. The first three focus on perinatal characteristics, including immediate postnatal growth, in relation to mammary gland mass and breast cancer risk, whereas the last two explore the relation of pregnancy hormones with breast cancer risk and with cellular populations that are likely to have mammary stem cell potential. The five projects are interlinked and they address the hypothesis that growth and mammotropic hormones in perinatal life affect the number of susceptible mammary gland cells. This number is likely to be reflected in birth size and rate of postnatal growth that, in turn, represent intermediate steps and correlates of mammary gland mass and breast cancer risk in adult life. The progress on each component project (CP) will be reported separately to facilitate the reader.

BODY

CP1 “Association of growth during the first postnatal week with breast cancer risk in adult life”

CP1 PI: Prof. Anders Ekbom, Unit of Clinical Epidemiology, Dept. of Medicine, Karolinska Institutet/Karolinska University Hospital, SE-171 76 Stockholm, Sweden.

Timetable of research accomplishments of CP1 as outlined in the Statement of Work.

Task 1  To investigate the association of growth during the first postnatal week with breast cancer risk in adult life:

a. Retrieval of available birth records from 1,068 women with incident breast cancer and 2,727 control women. (Months 1-24)

b. Extraction of data on growth of newborns during the first postnatal week, as well as information on covariates to be used in the analysis. (Months 25-30)

c. Linkage of data on postnatal growth and perinatal covariates to cancer and mortality registries. (Months 31-36)

d. Data analyses. (Months 37-48)

e. Manuscript preparation and submission. (Months 49-60)

CP1 progress report

The retrieval of birth records (task 1a) is ongoing. We foresee no problems in completing the retrieval by June 2007, within 24 months of the formal launch of CP1 (component project 1 was officially launched July 1, 2005).
The extraction of data on growth of newborns (task 1b) is ongoing. In the retrieval process, the extraction of data on growth is done in parallel in newborns through the first postnatal week. We foresee no problems to finalize this work within the proposed time schedule. We have so far encountered no problems in identifying and retrieving information as given in the research plan.

In accordance with the time plan, tasks 1c, d and e have not yet been initiated.

**CP1 key research accomplishments**

- The retrieval of birth records from 1,068 women with incident breast cancer and 2,727 control women is ongoing. So far, about 80% of the retrieval has been completed.
- The extraction of data on growth of newborns during the first postnatal week is ongoing. So far, about 80% of the retrieval has been completed.

**CP1 reportable outcomes**

No outcomes can be reported at this stage.

**CP1 Conclusion**

No major problems have been encountered and none are foreseen for the continuation of the implementation of the component project. As planned, retrieval of birth charts will be completed within 24 months from official launch of CP1 (July 1, 2005) and extraction of data on growth of newborns within 30 months from official launch of CP1.

---

**CP2: “Relation of perinatal characteristics and postnatal growth velocity with mammographic patterns in adult life”**

*CP2 PI: Prof. Per Hall, Dept. of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, SE-171 77 Stockholm, Sweden*

**Timetable of research accomplishments of CP2 as outlined in the Statement of Work.**

**Task 2** To investigate the relation of perinatal characteristics and postnatal growth velocity with mammographic patterns in adult life:

a. Retrieval of available birth records from 3,345 women with invasive breast cancer and 3,454 controls (these women are not the same with those to be studied in the context of task 1). (Months 1-48)

b. Retrieval of the available sequential mammographies of the women with breast cancer and the control women. (Months 1-48)

c. Linkage of mammographic data to perinatal characteristics and postnatal growth velocity. (Months 30-48)
d. Evaluation of mammographies through a computer-assisted grey-scale thresholding methods technique. (Months 25-36)

e. Data analyses. (Months 37-48)

f. Manuscript preparation and submission. (Months 49-60)

CP2 progress report

The retrieval of birth records (task 2a) is ongoing. Based on the experience gained during the first year of the project, we have estimated that of all records, about 60% will be retrievable and sufficiently informative. This is mostly due to the fact that for a large number of cases and control women, who were born in the 1920-30s, the records are not of sufficient informativeness and quality. So far, birth records for 1,200 women are retrieved.

The retrieval of available mammograms (task 2b) is ongoing. Based on the experience gained during the first year of the project, and due to the fact that some centers have destroyed their mammography films, we have estimated that we will be able to retrieve the mammograms of 60% of all women. Again, the problem is mainly concerning mammograms of women born in the 1920-30s. So far, we have found mammograms for 1,700 women (cases and controls), with on average three examinations before diagnosis per woman. For all women, we have been able to get the original films, so we do not have the anticipated problems with reduced quality of the copies.

Initially we planned to use the scanner at the Karolinska Hospital, in collaboration with Drs. Gunilla Svane and Eduardo Azavedo. As the capacity and quality of that scanner was not quite what we needed, we decided to use another scanner located at the Department of Medical Epidemiology and Biostatistics. We have recently started to scan the original mammography films, and as this scanner has the possibility to load with 100 films at a time, the process will be efficient. The scanner is a Fuji/Array 2905 with optical density 4.7, which exceeds the quality needs for this study. We have initiated collaboration with Isabel dos Santos Silva and Julian Peto and their group at London School of Hygiene and Tropical Medicine to assure good quality for the scanning and density reading methods. This group has since long studied the influence of mammographic density on breast cancer risk. The group has a vast experience in digitizing and evaluating density. As a validation study we have sent mammograms for 160 patients to London. Films will be scanned, evaluated and compared with our results.

The main responsible researcher will go from full-time to part-time but a junior researcher (graduate student) have been hired from February to work full-time (supervised by the responsible researcher).

The evaluation of mammographies through a computer-assisted grey-scale thresholding methods technique (task 2d) will be done with the Cumulus software, a semi-automatic computer assisted technique where the evaluator indicates the area to be measured. The program is installed, and the responsible Swedish scientists will take a course on the method given by the developers of the program, Norman Boyd and Martin Yaffe in Toronto in April 2007. The density readings will thus start by month 22 and is estimated to take 18 months. Given the large number of films
we have also participated, together with Julian Peto’s group, in a project aiming for a fully automatic measuring system developed by a group in Oxford.

In accordance with the timetable of CP2, task 2c,e have not yet been initiated.

CP2 key research accomplishments

- Birth records have so far been retrieved for 1,200 women (cases and controls).
  Mammograms have so far been retrieved for 1,700 women (cases and controls).
  Mammograms and birth records are expected to be retrieved for approximately 4,000 women.
- For patients with breast cancer, we now have detailed information on where and when they had their mammograms taken. For controls, mammography screening units have been identified and collection of films initiated.
- The film scanning process is ongoing and running smoothly.

CP2 reportable outcomes

At this stage, there are no reportable outcomes.

CP2 Conclusion

As outlined under “Research accomplishments”, we are on the way with the goals for task 2. We are up and running and have the necessary qualifications and manpower. The rate of retrieving information on birth characteristics and mammographic films is in line with what we anticipated and based on our experience from the first year of the project we anticipate to be able to retrieve birth records and mammograms for 60% of the women indicated in the CP2 proposal outline, i.e. for approximately 4,000 women. The collaboration with Professor Peto’s group has strengthened our competence and we are participating in a project aiming for a fully automatic measurement system.

No major problems have been encountered and none are foreseen for the continuation of the implementation of the component project.

CP3: “Interaction of perinatal characteristics with genes that are likely related to breast cancer risk”

CP3 PI: Prof. Per Hall, Dept. of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, SE-171 77 Stockholm, Sweden

Timetable of research accomplishments of CP3 as outlined in the Statement of Work.
Task 3 To investigate the possible interaction of perinatal characteristics with genes that are likely related to breast cancer risk:

a. Identification and selection of genes likely to be related to breast cancer risk, e.g. ESR1, AIB1, and the IGF family (Months 1-12)

b. Selection of "tagging" single nucleotide polymorphisms (tSNPs). The choice of tSNBs aims at avoiding redundant genotyping. A good marker coverage is expected to be achieved by using approximately one SNP per 3 Kb. (Months 1-12).

c. Genotyping of the approximately 8 genes selected for the study (Months 13-36).

d. Data analyses. (Months 36-48)

e. Manuscript preparation and submission. (Months 49-60)

CP3 progress report

A high percentage of dense parenchyma on mammographic images, which appears to confer a 4-to 6-fold increased risk for breast cancer, has a genetic component, based on the significant correlations with breast density observed between sisters and the 2-fold higher correlation between monozygotic compared to dizygotic twins. A role of steroid hormones in mammographic density is supported by observations of an increase in density after HRT and a decrease after suppression of ovarian function through a gonadotropin-releasing hormone agonist or after tamoxifen treatment. Genetically determined differences in biosynthesis and metabolic pathways of estrogens may affect breast cancer risk as reflected in mammographic density. We focused on the association of polymorphisms in ESR1, AIB1 and the IGF family.

DNA has been isolated for approximately 50% of the 2,818 cases and 3,111 controls who fulfilled all the criteria for participation in this component project. The DNA has been shipped to Genome Institute of Singapore where the SNP analyses will be performed. The genes that we will focus on are ESR1, ATM, IGF-1, CHEK2, HER-2. The CHEK2 gene was analyzed in a straight-forward association study (not part of the current US Army funded project since genetic variation was not correlated to mammographic density). This will be done in the future, when we have the mammography information.

With respect to task 3b, we are currently examining the ESR1 and have genotyped 155 SNPs in the gene, including 52 tSNPs. Results should be produced by late March 2007 and a draft manuscript is expected to be ready by late April 2007.

Analyses of the ESR1 are ongoing. Since the application was written the number of genetic association studies has exploded. Genetic alterations previously not considered related to breast cancer have been found to be associated to the disease. As a consequence of what we have learnt when participating in a large scale genetic association breast cancer consortium headed by Bruce Ponders group in Cambridge we want to study some additional genes. Among those that we would like to add are Casp8 that we just published in Nature Genetics (Cox et al, 2007) and CYP19A1 where we have initial and very promising results.

Genotyping (task 3c) has already been finished for the CHEK2 gene and will start shortly for the other genes mentioned under task 3a.
In accordance with the time table, tasks 3d-e have not yet been initiated.

**CP3 key research accomplishments**

- Work on *ESR1* is on going and a draft manuscript is expected to be ready by late April 2007.
- Preliminary results from other studies indicate that a variant of *CYP19A1* is a promising determinant of breast cancer risk. This gene and Casp8 will be studied in addition to the genes indicated in the initial proposal.

**CP3 reportable outcomes**

At this stage, there are no reportable outcomes.

**CP3 Conclusion**

As outlined under “Research accomplishments”, we are well on the way with the goals for task 3. We have the necessary qualifications and manpower. We foresee no problems to complete the different tasks within the calculated timeframe. Given the wealth of new information from genetic association studies and our own results, we will also study the role of a variant of *CYP19A1* and Casp8.

**CP3 References**


---

**CP4: “Pregnancy hormones and perinatal breast cancer risk factors in Boston, USA and Shanghai, China”**

**CP4 co-PIs:** Prof. Dimitrios Trichopoulos and Dr. Pagona Lagiou, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115,

Timetable of research accomplishments of CP4 as outlined in the Statement of Work.

**Task 4** To study maternal and cord blood levels of components of the IGF system and adiponectin among Caucasian women in North America and Chinese women in Asia in conjunction to maternal anthropometry and birth size parameters:

a. Retrieval of the available stored cord blood and maternal serum samples from 304 pregnant Caucasian women in Boston, US and 335 pregnant Chinese women in Shanghai, China and transfer of these samples to the laboratory for hormone determination. (Months 1-6)
b. Conduct of laboratory assays for hormones (Months 7-24)
c. Linkage of maternal and newborn data to maternal and cord blood hormone levels (Months 24-30)
d. Data analyses (each of the measured hormones in maternal and cord blood to be studied in conjunction to maternal and newborn variables) (Months 31-48)
e. Manuscript preparation and submission. (Months 37-60)

**CP4 progress report**

With respect to baseline data, the database is already in place, as also indicated in the 1st annual report. A total of 1,376 blinded maternal serum samples have been retrieved from among the stored samples, recorded and submitted to the UMass Medical School, ILAT Steroid RIA Laboratory, for hormonal assays. A total of 245 cord blood samples have also been retrieved from among the stored samples, recorded and submitted to the UMass Medical School, ILAT Steroid RIA Laboratory, for hormonal assays. As the samples are identified only with a serum sample number, not the study ID, they include some that were ineligible or with information missing on relevant study variables. These will be verified at the data merging stage before statistical analyses.

Tasks 4a and 4b have been completed. The laboratory has assayed testosterone, IGF-1, IGFBP-3, and adiponectin on all the maternal samples. The laboratory has also assayed these hormones plus estradiol, estriol, progesterone, SHBG and IGF-2 on the cord blood samples.

Task 4c has been initiated. In accordance with the timetable, tasks 4d and 4e have not yet been initiated.

**CP4 key research accomplishments**

- Determination of testosterone, IGF-1, IGFBP-3 and adiponectin in maternal serum samples has been completed.
- Determination of estradiol, estriol, progesterone, SHBG, testosterone, IGF-1, IGF-2, IGFBP-3 and adiponectin in cord blood samples has been completed.

**CP4 reportable outcomes**

At this stage, there are no reportable outcomes.

**CP4 Conclusion**

No major problems have been encountered and none are foreseen for the continuation of the implementation of the component project.
CP5 PI: Prof. Chung-Cheng Hsieh, University of Massachusetts Cancer Center, 55 Lake Avenue North, Worcester, MA 01655

Timetable of research accomplishments of CP5 as outlined in the Statement of Work.

Task 5 To investigate whether markers of mammary stem cells are associated with perinatal characteristics that are linked to breast cancer in later life:

a. Finalization of questionnaire for obtaining maternal and gestation characteristics. (Months 1-3)
b. Training of the study personnel on study procedures. (Months 1-6)
c. Subject recruitment and sample collection from a total of 250 pregnant women. (Months 7-42)
d. Conduct of laboratory assays for markers of stem cells (Months 7-45)
e. Conduct of laboratory assays for hormones (Months 10-48)
f. Data analyses. (Months 45-54)
g. Manuscript preparation and submission. (Months 49-60)

CP5 progress report

Clearance of component project 5 (a+b) was granted by the US Army Human Subjects Research Review Board (HSRRB) on July 17, 2006. Component project 5 was officially launched on July 18, 2006.

The underlying premises of this project are that 1) breast stem cells are the cell type that undergoes malignant transformation, 2) breast stem cells primarily arise during the fetal/perinatal period, and therefore the in utero/perinatal environment is a major determinant of the breast stem cell number in an individual, and 3) the greater the number of breast stem cells, the greater the likelihood of that one will undergo a genetic alteration that will be oncogenic. Previously we have shown that the concentration of hematopoietic stem cells in cord blood, serving as a surrogate for general stem cell potential, can be correlated to both perinatal levels of mitogens including estrogens and IGF-1, and to perinatal parameters such as birth weight, which have been linked to elevated breast cancer risk (Baik et al, 2004 and 2005). Ideally, one would like to obtain some indicators of the levels of epithelial precursors cells in the perinatal environment, and determine if such cells might be an even better indicator of future breast cancer risk: this is the goal of Component Project 5, using umbilical cord blood as the perinatal cell source.

Since the launching of the project on July 18, 2006, the questionnaire for obtaining maternal and gestation characteristics has been finalized and study personnel have been trained on the study procedures (Task 5a and 5b).
Subject recruitment and sample collection started at the Tuft-New England Medical Center on November 16, 2006 (Task 5c). From November 16, 2006 to March 15, 2007, 46 women consented to participate in the study. Of these 46 women, umbilical cord blood was collected from 31 women. Among the 31 women, 27 collections were delivered to the stem cell lab within the requisite 24 hours. We had complete participation (i.e. completion of questionnaire) among 26 of the 27 women for whom we collected cord blood and had sent to the LINK laboratory at the University of Massachusetts for analysis within 24 hours of delivery and collection.

In the past grant year, we have made progress toward developing a protocol for obtaining cells with epithelial markers from umbilical cord samples, and we are currently investigating whether these cells have stem cell-like markers and phenotypes (Task 5d). In this protocol, mononuclear cells (MNCs) are enriched from umbilical cord samples and then are immunomagnetically selected for expression of the epithelial cell-specific surface marker, EpCAM (also known as epithelial-specific antigen). EpCAM<sup>+</sup> cells have been obtained from all 24 of the umbilical cord samples processed for this marker, and represent anywhere from 1-13% of the MNCs; the viability of these cells after immunomagnetic sorting are generally >95%. FACS analysis on selected samples has demonstrated that >60% of isolated EpCAM<sup>+</sup> cells express cytokeratins, a definitive marker of epithelial cell types. EpCAM<sup>+</sup> cells are plated in a defined, serum-free medium consisting of MEGM (mammary epithelial cell growth medium; Cambrex) supplemented with B27, EGF, bFGF, and heparin; this is referred to as “Mammosphere” medium. When plated on regular, uncoated plastic dishes or on dishes coated with collagen IV or fibronectin, the cells attach to the substrate and remain quiescent. However, when plated on dishes coated with collagen I, the cells remain in suspension and within one week’s time form spherical clusters. This finding is potentially significant as nonadherent cluster formation in a serum-free environment has been observed as a phenotype of primitive, stem-like cells from both neural and mammary tissues. EpCAM<sup>+</sup> cells generally do not form spheres in culture. Although these clusters can be disrupted and passaged up to 4 times, the cells eventually senesce. The number of “spheres” or clusters can be quantitated using a low-power inverted scope and a grid. In preliminary results, the number of clusters range from 0.03 to 0.34% of the number of plated EpCAM<sup>+</sup> cells. Future experiments will determine 1) the potential of these EpCAM<sup>+</sup> cells for clonal growth, 2) the expression of embryonic markers e.g. Oct-4 and Nanog in these cultured cells, 3) the expression of breast stem cell associated markers, included nestin, CD44, CD24, and CD49f in lineage-negative cells 4) the ability of these cells to differentiate into mammary tissue cell subtypes, including luminal cells (e.g. cytokeratin 18, sialomucin-positive) or myoepithelial cells (cytokeratin 14, α-smooth muscle actin-positive), and 5) whether supplementation with bovine pituitary extract, HGF, or other factors will abrogate cell senescence, and maintain any stem-like phenotypes.

CP5 key research accomplishments

- The questionnaire for obtaining maternal and gestation characteristics has been finalized.
- Training of the study personnel on study procedures has been completed.
- Subject recruitment and sample collection have started.
- Laboratory assays for markers of stem cells are being conducted
CP5 reportable outcomes

At this stage, there are no reportable outcomes.

CP5 Conclusion

Although the component project 5 had a delayed launching, no major problems are foreseen for its implementation.

CP5 References


Task 6: “Monitoring, coordination and fine-tuning of the five component projects”

Monitoring and coordination of the five component projects has presented no problems. The key investigators have a long history of successful scientific collaboration, which continues in the context of the current project.

KEY RESEARCH ACCOMPLISHMENTS

- The retrieval of birth records from 1,068 women with incident breast cancer and 2,727 control women is ongoing (CP1).
- The extraction of data on growth of newborns during the first postnatal week is ongoing (CP1).
- Birth records have so far been retrieved for 1,200 women (cases and controls). Mammograms have so far been retrieved for 1,700 women (cases and controls). Mammograms and birth records are expected to be retrieved for approximately 4,000 women (CP2).
- For patients with breast cancer, we now have detailed information on where and when they had their mammograms taken. For controls, mammography screening units have been identified and collection of films initiated (CP2).
- The film scanning process is ongoing and running smoothly (CP2).
- Work on $ESRI$ is on going and a draft manuscript is expected to be ready by late April 2007 (CP3).
• Preliminary results from other studies indicate that a variant of \textit{CYP19A1} is a promising determinant of breast cancer risk. This gene and Casp8 will be studied in addition to the genes indicated in the initial proposal (CP3).
• Determination of testosterone, IGF-1, IGFBP-3 and adiponectin in maternal serum samples has been completed (CP4).
• Determination of estradiol, estriol, progesterone, prolactin, SHBG, testosterone, IGF-1, IGF-2, IGFBP-3 and adiponectin in cord blood samples has been completed (CP4).
• The questionnaire for obtaining maternal and gestation characteristics has been finalized (CP5).
• Training of the study personnel on study procedures has been completed (CP5).
• Subject recruitment and sample collection have started (CP5).
• Laboratory assays for markers of stem cells are being conducted (CP5).

\textbf{REPORTABLE OUTCOMES}

At this stage, there are no reportable outcomes.

\textbf{CONCLUSIONS}

Overall no major problems have been encountered and none are foreseen for the continuation of the implementation of the project.

\textbf{REFERENCES}


\textbf{APPENDICES}

-