



**Karolinska  
Institutet**

**Institutionen för medicinsk epidemiologi och biostatistik**

# Genetic Association and Risk Prediction of Breast Cancer from an Epidemiological and Biostatistical Perspective

**AKADEMISK AVHANDLING**

som för avläggande av medicine doktorsexamen vid Karolinska  
Institutet offentligen försvaras i Petrén, Nobels väg 12B

**Fredagen den 9 mars, 2012, kl 09.00**

av

**Hatef Darabi**

*Huvudhandledare:*

Docent Keith Humphreys  
Karolinska Institutet  
Institutionen för medicinsk epidemiologi och  
biostatistik

*Bihandledare:*

Professor Kamila Czene  
Karolinska Institutet  
Institutionen för medicinsk epidemiologi och  
biostatistik

Professor Per Hall  
Karolinska Institutet  
Institutionen för medicinsk epidemiologi och  
biostatistik

*Fakultetsopponent:*

Dr Antonis Antoniou  
University of Cambridge  
School of Clinical Medicine

*Betygsnämnd:*

Professor Matteo Bottai  
Karolinska Institutet  
Institutionen för miljömedicin

Docent Helena Jernström  
Lund Universitet  
Institutionen för kliniska vetenskaper

Docent Johan Askling  
Karolinska Institutet  
Institutionen för medicin

**Stockholm 2012**

## ABSTRACT

The main purposes of this thesis were to analyse common genetic variation in candidate genes and candidate pathways in relation to breast cancer risk, prognosticators and survival, to develop statistical methods for genetic association analysis for evaluating the joint importance of genes, and to investigate the potential impact of adding genetic information to clinical risk factors for projecting individualised risk of developing breast cancer over specific time periods.

In Paper I we studied genetic variation in the estrogen receptor  $\alpha$  and epidermal growth factor genes in relation to breast cancer risk and survival. We located a region in the estrogen receptor  $\alpha$  gene which showed a moderate signal for association with breast cancer risk but were unable to link common variation in the epidermal growth factor gene with breast cancer aetiology or prognosis.

In Paper II we investigated whether suspected breast cancer risk SNPs within genes involved in androgen-to-estrogen conversion are associated with breast cancer prognosticators grade, lymph node status and tumour size. The strongest association was observed for a marker within the CYP19A1 gene with histological grade. We also found evidence that a second marker from the same gene is associated with histological grade and tumour size.

In Paper III we developed a novel test of association which incorporates multivariate measures of categorical and continuous heterogeneity. In this work we described both a single-SNP and a global multi-SNP test and used simulated data to demonstrate the power of the tests when genetic effects differ across disease subtypes.

In Paper IV we assessed the extent to which recently associated genetic risk variants improve breast cancer risk-assessment models. We investigated empirically the performance of eighteen breast cancer risk SNPs together with mammographic density and clinical risk factors in predicting absolute risk of breast cancer. We also examined the usefulness of various prediction models considered at a population level for a variety of individualised breast cancer screening approaches.

The goal of a genetic association study is to establish statistical associations between genetic variants and disease states. Each variant linked to a disease can lead the way to a better understanding of the underlying biological mechanisms that govern the development of a disease. Increased knowledge of molecular variation provides the opportunity to stratify populations according to genetic makeup, which in turn has the potential to lead to improved disease prevention programs and improved patient care.

ISBN 978-91-7457-667-2