THE ROLE OF THE -160 C/A SINGLE NUCLEOTIDE POLYMORPHISM IN THE CDH1 GENE AND E-CADHERIN EXPRESSION IN ETIOLOGY, PROGRESSION, AND SURVIVAL IN INVASIVE BREAST CANCER

by

Sandra L. Deming

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology, School of Public Health.

Chapel Hill
2005

Approved by:

Advisor: Professor Robert C. Millikan

Reader: Professor Andrew F. Olshan

Reader: Professor Jane C. Schroeder

Reader: Professor Michael Symons

Reader: Professor Leigh Thorne

TABLE OF CONTENTS

Page

LIST OF TABLES.....xi

LIST OF FIGURES.....xv

LIST OF ABBREVIATIONS.....xvi

CHAPTER

I. BACKGROUND

Introduction.....1

Literature Review.....1

Public health significance.....1

Breast cancer heterogeneity and classification.....2

Risk factors for breast cancer.....5

Prognostic factors for breast cancer.....17

E-cadherin: Structure and general description.....21

E-cadherin in the estrogen/breast cancer hypothesis.....22

-160 C.A SNP in the e-cadherin gene.....24

E-cadherin expression in breast tissue and breast cancer.....28

References.....38

Tables and figures.....56

II. STUDY DESIGN AND METHODS....57

Summary of study....58

Specific aims....60

Materials and methods....63

Description of study population: The Carolina Breast Cancer Study....64

Data collection....65

Laboratory methods....67

Data analysis....71

Power calculations....80

Reliability results....86

References....88

Tables and figures....101

III. RESULTS: E-CADHERIN EXPRESSION, CDH1 GENOTYPE AND RISK OF BREAST CANCER....105

Statistical analysis....105

Evaluating potential selection bias....107

E-cadherin and risk of breast cancer....110

CDH1 polymorphism frequencies....112

ORs for breast cancer and *CDH1* genotype....113

CDH1 genotype and e-cadherin tumor expression....114

References....114

Tables and figures....118

IV. RESULTS: E-CADHERIN EXPRESSION, CDH1 GENOTYPE AND SURVIVAL....144

Determination of death.....144

Survival time.....145

Censoring status.....145

Kaplan Meier Survival Analysis.....146

Survival over entire period of observation.....146

Five year survival.....146

Demographic and tumor characteristics and survival.....147

E-cadherin protein expression and breast cancer survival.....149

CDH1 -160 SNP and breast cancer survival.....150

Missing Values.....151

Cox proportional hazards analysis.....152

All-cause mortality vs. breast cancer mortality..... 152

Evaluation of proportional hazards assumption.....152

Identification of potential confounders.....154

Evaluation of potential covariate interactions with e-cadherin Protein expression and *CDH1* genotype.....156

Independent prognostic value of e-cadherin protein expression in breast cancer mortality.....157

E-cadherin and lymph node status.....158

Independent prognostic value of the CDH1 -160 C/A genotype in breast cancer mortality.....158

References.....160

Tables and figures.....161

V. DISCUSSION AND SUMMARY.....194

E-cadherin expression, *CDH1* genotype, and risk of breast cancer.....194

Risk factors and e-cadherin defined breast cancer.....194

CDH1 genotype and e-cadherin tumor expression.....196

CDH1 genotype and risk of breast cancer.....198

Summary of e-cadherin expression, *CDH1* genotype and risk of breast cancer.....200

E-cadherin expression, *CDH1* genotype, and survival.....202

e-cadherin expression and survival.....202

CDH1 genotype and survival.....202

Survival summary discussion.....204

Public health significance.....207

Future directions.....208

Conclusions.....210

References.....212

LIST OF TABLES

Table 2.1 Response rates of CBCS (Phase I): IHC portion of study.....102

Table 2.2 Response rate of CBCS cases and controls (Phase I, II and CIS); Genotype portion of study.....103

Table 2.3 Tumor block retrieval rates.....104

Table 2.4 Power to detect and association among the different subgroups given various prevalence of exposures.....105

Table 2.5 Power to detect a gene environment interaction.....106

Table 2.6 Power to detect a gene environment interaction.....107

Table 2.7a E-cadherin intra-rater reliability scoring results: Overall e-cadherin positivity.....108

Table 2.7b E-cadherin intra-rater reliability scoring results: Intensity.....108

Table 2.7c E-cadherin intra-rater reliability scoring results: Percent cells positive.....108

Table 2.7d E-cadherin intra-rater reliability scoring results: Cellular localization.....109

Table 2.7e E-cadherin intra-rater reliability scoring results: Distribution.....109

Table 2.7f E-cadherin intra-rater reliability scoring results: Pattern.....109

Table 2.8 E-cadherin tumor positivity by scorer.....110

Table 2.9a Overall e-cadherin positivity by scorer: (SD vs. DC).....111

Table 2.9b Overall e-cadherin positivity by scorer: (SD vs. LT).....111

Table 2.9c Overall e-cadherin positivity by scorer: (DC vs. LT).....111

Table 2.10a Intensity of e-cadherin staining by scorer: (SD vs. DC).....112

Table 2.10b Intensity of e-cadherin staining by scorer: (SD vs. LT).....112

Table 2.10c Intensity of e-cadherin staining by scorer: (DC vs. LT).....112

Table 2.11a Percent of cells positive by scorer: (SD vs. DC).....113

Table 2.11b Percent of cells positive by scorer: (SD vs. LT).....113

Table 2.11c Percent of cells positive by scorer: (DC vs. LT).....113

Table 2.12 Inter-rater reliability of other e-cadherin scoring parameters.....114

Table 2.13 Summary of tumor sections available for staining.....115

Table 3.1 Comparison of characteristics of cases with and without e-cadherin protein expression information.....132

Table 3.2 Distribution of *CDH1* genotype by various demographic characteristics among CBCS phase I and II controls.....135

Table 3.3 Comparison of characteristics of cases with and without *CDH1* genotype information among CBCS phase I and II cases.....137

Table 3.4 Odds ratios for breast cancer in relation to various known and Suspected risk factors in CBCS.....140

Table 3.5 Distribution of e-cadherin expression by various demographic and tumor characteristics among CBCS participants.....142

Table 3.6 Odds ratios and 95% confidence intervals for e-cadherin+ and e-cadherin- controls in CBCS.....145

Table 3.7 Odds ratios and 95% confidence intervals for e-cadherin + and e-cadherin- versus controls in CBCS individually adjusted for ER status, PR status, or stage.....149

Table 3.8 *CHD1* allele frequencies among cases and controls with *CDH1* genotype Information.....151

Table 3.9 Distributions and odds ratios for *CDH1* genotype among cases and controls with *CDH1* genotype information and test for Hardy Weinberg information.....152

Table 3.10 Odds ratios for breast cancer cases versus controls by *CDH1* genotype.....153

Table 3.11a P-values for potential interactions between *CDH1* genotype and various estrogen-related covariates in overall odds of breast cancer.....154

Table 3.11b Odds ratios for potential interactions between *CDH1* genotype and oral contraceptive use, hormone replacement therapy use, and age at menarche in all breast cancer.....154

Table 3.12a Distribution of e-cadherin protein expression by *CDH1* genotype.....155

Table 3.12b Distributions of e-cadherin protein expression by *CDH1* genotype Based on 8-level categorization.....155

Table 3.13 Distributions and odds ratio for *CDH1* genotype and e-cadherin + and e-cadherin – breast cancer.....156

Table 3.14a P-values for potential interactions between *CDH1* genotype and various estrogen-related covariates in e-cadherin delineated breast cancer.....157

Table 3.14b P-values for potential interactions between *CDH1* genotype and oral Contraceptive use and parity in e-cadherin delineated breast cancer.....157

Table 4.1 Summary of classification of event and censoring date for various subgroups utilized in the survival analysis.....175

Table 4.2 Number of individuals having event or being censored in various analytic subgroups.....176

Table 4.3 5-year and entire period of observation all-cause survival by demographic characteristics in patients with invasive breast cancer.....177

Table 4.4a 5-year and entire period of observation breast cancer-specific survival by demographic characteristics in patients with invasive breast carcinoma.....178

Table 4.4b Joint effects of race and menopausal status on survival over entire period of observation.....179

Table 4.5 5-year and entire period of observation all-cause survival by tumor characteristics in patients with invasive breast carcinoma.....180

Table 4.6 5-year and entire period of observation breast cancer-specific survival by tumor characteristics in patients with invasive breast cancer.....182

Table 4.7a 5-year and entire period of observation all-cause survival based on e-cadherin status in various subgroups in patients with invasive breast carcinoma.....184

Table 4.8 5-year and entire period of observation breast cancer-specific survival based on e-cadherin status in various subtypes.....186

Table 4.8b Joint effect of e-cadherin expression and estrogen receptor status on survival over entire period of observation.....185

Table 4.9 5-year and entire period of observation all-cause survival based on *CDH1* genotype.....187

Table 4.10 5-year and entire period of observation breast cancer-specific survival based on *CDH1* genotype.....188

Table 4.11 Univariate hazard ratios and 95% confidence intervals for various Factors for the entire period of observation.....189

Table 4.12 Preliminary assessment of confounding with e-cadherin protein expression among breast cancer deaths (5-year follow-up).....191

Table 4.13 Preliminary assess of confounding with e-cadherin protein expression among breast cancer deaths (entire period of follow-up).....192

Table 4.14 Preliminary assessment of confounding with *CDH1* genotype among breast cancer deaths (5-year follow-up).....193

Table 4.15 Preliminary assessment of confounding with *CDH1* genotype among breast cancer deaths (entire period of observations).....194

Table 4.16 P-values for potential interactions between e-cadherin expression and various covariates.....195

Table 4.17 P-values for potential interactions between *CDH1* genotype and various covariates.....196

Table 4.18 5-year and entire period of observation Cox proportional hazard ratios for breast cancer deaths based on e-cadherin tumor expression.....197

Table 4.19 5-year and entire period of observation Cox proportional hazard ratios for breast cancer deaths based on e-cadherin tumor expression stratified by lymph node status.....198

Table 4.20 5-year and entire period of observation Cox proportional hazard ratios for breast cancer deaths based on *CDH1* genotype.....199

Table 4.21 5-year and entire period of observation Cox proportional hazard ratios for breast cancer deaths based on *CDH1* genotype stratified by P53 functional mutation status.....200

LIST OF FIGURES

Figure 1.1 E-cadherin and its relationship to other cellular proteins.....56

Figure 1.2 Model of factors and presumed mechanisms involved in the estrogen and e-cadherin pathway.....57

Figure 2.1 Examples of e-cadherin immunohistochemical staining in normal Breast tissue, ductal carcinoma *in situ,* and invasive lobular and invasive ductal breast cancer.....117

Figure 2.2 Scoring sheet: e-cadherin expressin in tumors of Phase I Carolina Breast Cancer Study.....118

Figure 4.1 5-year breast cancer-specific survival by e-cadherin tumor expression status.....201

Figure 4.2 Entire period of observation breast cancer-specific survival by e-cadherin tumor expression status.....202

Figure 4.3 5-year breast cancer-specific survival by *CDH1* genotype.....203

Figure 4.4 5-year breast cancer-specific survival by e-cadherin tumor expression status with missing data plotted.....204

Figure 4.5 5-year breast cancer-specific survival by *CDH1* genotype with missing data plotted.....205

Figure 4.6 log-log plot of months survived versus the survival distribution function for e-cadherin tumor expression to assess proportional hazards assumption.....206

Figure 4.7 log-log plot of months survived versus the survival distribution function for *CDH1* genotype to assess proportional hazards assumption....207