CAUTION: The following text, a synopsis of my dissertation, may make you drowsy. Do not read while operating heavy equipment, driving, or riding your bike. Best if read when you are suffering from insomnia or stuck on a long flight

Joanna, April 1, 2001

GENE-ENVIRONMENT INTERACTION IN BIRTH DEFECTS

By

Joanna Sue Zeiger

A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, MD
May, 2001

ABSTRACT

In the U.S., birth defects occur in 3-5% of live-births and comprise the leading cause of infant mortality, producing a substantial economic, social, and emotional burden. The case-control studies presented in this dissertation tested two major questions: 1) does an association exist between environmental factors, candidate genes and birth defects?; and 2) are there interactions between these genes and environmental factors? Article 1, *Testing for Gene-Environment Interaction and Oral Clefts using Two Control Groups*, combined case groups from two studies and compared them to two separate control groups: (1) non-cleft malformations and (2) non-malformed infants. Genotype at the TGFA locus did not increase risk for oral clefts. There was an increased risk for cleft lip and palate (CL/P) and cleft palate only (CP) when compared to the non-malformed controls if the mother reported smoking during pregnancy. There was also evidence for gene-environment interaction among the CP group (malformed controls: OR=3.47, 95% CI=1.51-7.88; non-malformed controls: OR=4.21, 95% CI=1.48-12.79). In Research Study 2, *Genetic and Environmental Risk Factors for Sagittal Craniosynostosis* there were no associations with any of the examined environmental factors. All cases that were sequenced were <u>negative</u> for mutations at the following candidate genes: FGFR1 exon IIIa, FGFR2 (exons IIIa and IIIc,), FGFR3 exon IIIa , and TWIST exon I. In the third study, *Methylenetetrahydrofolate Reductase and Congenital Heart Defects: Evidence for Gene-Environment Interaction*, congenital heart defect cases and controls were genotyped for two MTHFR variants, C677T and A1298C. There were no differences between cases and controls for either of these MTHFR variants. Heavy alcohol use (OR=6.76, 95% CI=2.61-17.51) and

heavy smoking (OR=4.70, 95% CI=1.38-16.04) increased risk for pulmonic stenosis among Caucasians. There was evidence of gene-environment interaction among the African-American pulmonic stenosis cases with maternal smoking and the MTHFR A1298C C- genotype (nonsmokers: OR=0.98, 95%CI=0.27-3.42; smokers: OR=3.03, 95%CI=1.09-8.52). These studies strengthen our understanding of the complex interaction between genes and environment in the etiology of birth defects.

BACKGROUND AND SIGNIFICANCE

Birth defects represent an important public health issue, in the United States birth defects comprise the leading cause of infant mortality (20%) and years of potential life lost before age 65 (Yang et al., 1997). No major changes in frequency of congenital anomalies have occurred in the last 50 years; however, there is a declining trend in mortality rates from birth defects due to improved medical care and increased use of prenatal diagnosis. Congenital anomalies occur in about three to five percent of all live births. Defects range from non-life threatening malformations such as extra digits to fatal anomalies such as anencephaly.

Many different factors play a role in the etiology of congenital malformations: genetic causes, 28%; multifactorial inheritance, 23%; teratogens, 3%; unknown, 43% (estimated by Nelson and Holmes, 1989). It is evident that birth defects are heterogeneous in their etiology; with an estimated 23% of birth defects considered multifactorial in nature, and the challenge is to determine the genetic and environmental factors that are involved. Genetic linkage studies, association studies, and family studies have shown associations between birth defects and oncogenes (e.g. *BCL3* and oral clefts), growth factor genes (*TGFA* and oral clefts), and homeobox genes (*CSX* and

heart defects). Environmental exposures that have been implicated in causing birth defects include maternal infections [rubella and cardiovascular defects, CNS and hearing defects, use of alcohol [oral clefts, cigarettes [oral clefts, craniosynostosis, heart defects, and drug use [antiepileptics and cardiovascular defects, oral clefts, and neural tube defects. Other studies have shown that genetic factors may interact with environmental factors to produce birth defects [smoking, TGFA and oral clefts.

The last several years have seen advances in both molecular genetics and statistical methods. This has enabled researchers to more adequately assess not only environmental risk factors and their association with disease, but also genetic risk factors and how genes and environment may work together to increase risk to disease. Several study designs exist to examine the combined effects of genes and environment, ranging from traditional case-control study designs which incorporates standard epidemiologic data as well as marker information on candidate genes to family based study designs.

The case-control study design has been commonly utilized to measure gene-environment interaction in birth defects, as the studied outcomes are rare, and multiple exposures can be assessed. Two conditions must be met: (1) the environmental exposure is independent of the genotype, and (2) misclassification of the environmental exposure is non-differential with respect to the genotype.

Appropriate selection of cases and controls is paramount to the case-control study design. When selecting cases, a clear definition of the case must first be devised, and then methods of identifying the cases must be developed. Improper or inconsistent diagnosis of cases can lead to misclassification and ultimately bias the results. In

addition, different studies may devise their own case definition, which could lead to conflicting results and an apparent lack of reproducibility.

Selecting a suitable control group in studies of gene-environment interaction is crucial; an inappropriate control group could lead to spurious findings as a result of confounding due to population stratification. For this reason, it is important that controls have a similar genetic background to cases, and when doing studies of candidate genes, analyses should be stratified by race. Population based studies provide an effective means of obtaining a representative sample of cases and controls.

Several limitations exist when performing case-control studies. Participation rates may be low and often differ between cases and controls. There is also the potential for both differential and nondifferential misclassification. With nondifferential misclassification, both cases and controls make errors in exposure recall; there is increased potential for this to occur if the time between the exposure and the interview is prolonged. Differential misclassification, also referred to as recall bias, is said to occur when cases remember past exposures more thoroughly than controls causing an apparent effect.

Many factors, both genetic and environmental, contribute to the occurrence of most, if not all, congenital malformations. This has created considerable obstacles in understanding their pathogenesis. Studying candidate genes elected for their biological relevance to the disease and identifying environmental risk factors may provide new information to understand the etiology of these birth defects and hopefully open up opportunities for effective strategies for primary prevention and genetic counseling.

Genetic and Environmental Risk Factors for Sagittal Craniosynostosis

Zeiger J, MS; Beaty T, PhD; Hetmanski J, MS; Wang H, MD, PhD; et al.

ABSTRACT

Objective: To investigate whether genetic and environmental factors influence risk for sagittal craniosynostosis.

Design: Data were derived from a clinic based case-control study.

Setting: Cases were ascertained from craniofacial clinics in the Baltimore-Washington region. Controls were recruited from a pediatric practice in Baltimore County.

Patients: Fifty-one Caucasian probands with isolated, non-syndromic sagittal craniosynostosis born in the Mid-Atlantic region were included in this analysis. Controls are Caucasian infants born in Maryland without any known birth defect (n=109).

Main Outcome Measures: Odds ratios (OR) and corresponding 95% confidence intervals (CI) were calculated. Cases were genotyped at several loci implicated in malformation syndromes that include craniosynostosis.

Results: There were no elevated risks for craniosynostosis when considering maternal or paternal smoking and maternal vitamin usage. Case mothers consumed less alcohol (OR=0.10, 95% CI 0.04-0.26) and had less education than control mothers (p<0.001). All cases sequenced were <u>negative</u> for mutations at the following genes: FGFR1 exon IIIa 755C->G, FGFR2 (exons IIIa and IIIc,), FGFR3 exon IIIa, and TWIST exon I.

Conclusions: These findings suggest that while TWIST and the FGFR genes are important in syndromic craniosynostosis; they are unlikely to be involved in isolated sagittal craniosynostosis. Parental education and alcohol consumption were associated with sagittal craniosynostosis.

Testing for Gene-Environment Interaction and Oral Clefts using Two Control Groups

J. Zeiger, T. Beaty, J. Hetmanski, I. McIntosh, M.Mracna, S-J. Hwang, K-Y. Liang, S. Panny, and C. VanderKolk

ABSTRACT

Purpose: This case-control analysis examines the effects of maternal smoking and infant genotype at the *Taq1* polymorphism in the transforming growth factor alpha locus on risk to oral clefts using data from two published studies.

Methods: Cases were Caucasian infants with an isolated cleft born in Maryland. Two controls groups were used: infants with isolated non-cleft birth defects and healthy infants. Polytomous logistic regression analyses were performed to estimate odds ratios and their 95% confidence intervals. All individuals were genotyped for the *Taq1* polymorphism.

Results: Genotype at the transforming growth factor alpha locus did not increase risk for oral clefts. There was an increased risk for cleft lip and palate (CL/P) and cleft palate only (CP) when compared to the healthy controls if the mother reported smoking during pregnancy (CL/P: OR=2.38, 95% CI=1.16-4.89; CP: OR=3.05, 95% CI=1.61-5.79). There was also evidence for statistically significant gene-environment interaction among the CP group.

Conclusions: These findings suggest that maternal smoking during pregnancy may be a general risk factor for oral clefts independent of genotype and there may be etiology heterogeneity between types of clefts.

Methylenetetrahydrofolate Reductase and Congenital Heart Defects: Evidence for Gene-Environment Interaction

Zeiger J, Hetmanski J, McIntosh I, Mracna M, Dey S, Ochiai L, Loffredo C, Beaty TH

ABSTRACT

Congenital heart defects represent a major public health issue; collectively they are the most common birth defect creating a considerable medical, emotional and economic burden. This case-control study examined two common poylmorphisms, C677T and A1298C, in the gene encoding for methylenetetrahydrofolate reductase. We tested for their possible involvement in the etiology of four common types of congenital heart defects: aortic stenosis (n=36), coarctation of the aorta (n=64), pulmonic stenosis (n= 120), and atrial septal defect (n=108). Odds ratios with 95% confidence intervals were calculated to test for associations between genotype and case/control status, and to test for interactions between genotype and environmental factors. There were no differences between cases and controls (n=480) for either of the methylenetetrahydrofolate reductase variants. Heavy alcohol use (OR=6.76, 95% CI=2.61-17.51) and heavy smoking (OR=4.70, 95% CI = 1.38-16.04) increased risk for pulmonic stenosis among Caucasians. We did not find strong evidence of association with these genotypes with these specific types of CHD, but there was some evidence of interaction between maternal smoking and the 1298C genotype among African American pulmonic stenosis cases (OR nonsmokers: 0.98, 95% CI= 0.27-3.42; OR smokers: 3.03, 95% CI=1.09-8.52).

Gene-Environment Interaction and Risk to Oral Clefts

J.S. Zeiger and T.H. Beaty

INTRODUCTION

Oral clefts include cleft lip (CL), cleft palate (CP) and cleft lip and palate (CLP) and collectively these constitute a heterogeneous group of non-fatal birth defects known to be multifactorial in origin, in that both genes and environmental factors contribute to their etiology. It is possible to reduce heterogeneity in a sample of oral cleft cases by eliminating infants with recognized malformation syndromes (genetic or teratogenic) that can include oral clefts, and infants with multiple anomalies that may not fall into any recognized malformation syndrome. Even so isolated, non-syndromic oral clefts represent a complex and heterogeneous group of birth defects where there is strong evidence of an etiologic role for both genetic and environmental factors. It is imperative, therefore, to design studies of oral clefts to consider the effects of both genes and environmental factors, as well as their possible interaction. A variety of study designs are available, each with their own advantages and limitations. Here we review how these different study designs can be used to test for gene-environment interaction and then describe published studies of oral clefts that incorporate gene-environment interactions to one degree or another.

In epidemiologic studies, interaction occurs when the joint effect of two risk factors is greater (synergism) or less than a simple combination of their individual effects. In the context of gene-environment interaction (GxE), the isolated effects of a genotype (or of an allele) and of some observed environmental exposure fail to predict the effect seen when both are present. While interactions are commonly described as

synergistic effects, it is important to realize that a number of patterns of risk in the presence of the gene alone, environment alone, and both are possible, including negative or protective effects. Investigators have described five "biologically plausible" models of relationships between genotype, environmental exposure, and their effect on disease risk. These models include the following: (1) the genotype increases expression of the risk factor, but which can act on its own; (2) the high risk genotype exacerbates the effect of the risk factor, but the genotype has no effect in unexposed individuals; (3) the exposure exacerbates the effect of the high-risk genotype, but there is no effect in individuals who are exposed if they have the low-risk genotype; (4) both exposure and the high risk genotype are needed to alter risk; and (5) the exposure and the genotype both effect disease risk separately, but the risk is higher or lower when they occur simultaneously.

The marginal effects of the gene or the environment may or may not be apparent when interaction does exist. A number of epidemiologic study designs can be used to test for gene-environment interaction, as discussed below, and several of these have been used to demonstrate gene-environment interaction in studies of oral clefts.

CONCLUSIONS

Studying the causes of birth defects is difficult; both genetic and environmental factors contribute to risk, and the possibility that genes and environment act together to determine risk cannot be overlooked . To address this latter issue, gene-environment interaction has become a primary focus in the study of birth defects. While pharmacogenetic studies have led the field in demonstrating that different genotypes respond variously to environmental exposures, the field of birth defects research has not yet been able to identify biologic mechanisms by which exogenous factors (such as smoking or alcohol) affect gene function to increase risk to the developing fetus.

Several different study designs and statistical tests exist to assess the joint effects of genes and environment, with case-control studies being the traditional method by which gene-environment interaction can be assessed in the context of birth defects. Cases with the disease of interest are compared to unaffected controls. Both groups are genotyped for the markers of interest and environmental exposure information is collected. The ability to test for the marginal effects of genotype and environmental exposures are two advantages to this case-control study design; however, the ever present possibility of population stratification which can lead to spurious associations is perhaps the biggest drawback to this design.

Case-only and family-based study designs have been devised to help overcome the problems inherent in the case-control study design. Case-only analyses involve comparing cases without the susceptibility genotype (i.e. the control group) to cases with the susceptibility genotype with respect to the presence or absence of the environmental exposure. The main effects of either genotype or the environmental risk

factor cannot be tested separately, but the case-only design offers an inexpensive way to test for gene-environment interaction.

Family-based designs utilize sibs, parents or cousins to test for gene-environment interaction. In these studies, each case is compared to an unaffected relative who shares genes to some extent. Case-parent trio designs allow for a case/control analysis using parents to define a "matched" control as the allele (or genotype) not transmitted to the affected child, such as the transmission disequilibrium test (TDT). This type of analysis has the advantage of controlling for heterogeneity between populations which can lead to false associations, but requires DNA from the parents.

It is also important to realize that detecting an association between a genetic marker and a disease does not imply evidence of a biological mechanism, as the association could arise from: (1) the genetic marker playing a role in disease etiology, (2) an indirect association between the genetic marker and a susceptibility locus nearby, whereby the two are linked and associated with each other, and (3) confounding due to population stratification.

Summary of studies

The studies presented here are examples of different ways to test for gene-environment interaction in the context of association studies. These studies tested two major questions: 1) Is there an association between environmental factors, candidate genes and birth defects?; and 2) are there interactions between these genes and the observable environmental factors (e.g. maternal smoking or alcohol use)? Additionally,

the three original studies presented in this dissertation exemplify the difficulty in identifying risk factors for birth defects.

Research Article 1, **Testing for Gene-Environment Interaction and Oral Clefts using Two Control Groups**, combined case groups from two studies to form larger case groups and compared them to two separate control groups, one with non-cleft birth defects and the other a group of non-malformed infants. Combining the case groups from two sources (a passive state-wide registry and a prospective study) allowed for more power to assess the impact of the genotype, cigarette smoking and their interaction.

Findings

Results of this analysis confirmed earlier studies. Genotype at the transforming growth factor alpha (TGFA) locus did not increase risk for oral clefts, i.e. there was no marginal effect of genotype alone. There was an increased risk for cleft lip and palate (CL/P) and cleft palate only (CP) when compared to the healthy controls if the mother reported smoking during pregnancy (CL/P: OR=2.38, 95% CI=1.16-4.89; CP: OR=3.05, 95% CI=1.61-5.79), regardless of TGFA genotype. There was also evidence for statistically significant gene-environment interaction among the CP group (BDRIS: OR=3.47, 95% CI=1.51-7.88; MAOCS: OR=4.21, 95% CI=1.48-12.79). These findings suggest that maternal smoking during pregnancy may be a general risk factor for oral clefts independent of genotype and there may be etiology heterogeneity between types of clefts. While smoking appeared to increase risk to both CLP and CP, evidence of this gene-environment interaction between TGFA genotype and maternal smoking was

limited to the CP group, suggesting the possibility of etiologic heterogeneity between CP and CLP.

Future directions

In the oral cleft sample, cases and parents were genotyped at both MTHFR markers. Preliminary analyses did not show increased transmission of either variant, and thus far, there is no evidence of gene-environment interaction (Beaty et al., 2001). Further analyses will be performed to examine gene-gene interaction with other genetic markers that have been genotyped in this sample.

Research article 2, **Genetic and Environmental Risk Factors for Sagittal Craniosynostosis**, also used a case-control design. However, only the cases were genotyped leaving the case-only design as the only option to test for gene-environment interaction in this sample.

Findings

There were no associations with any of the available environmental factors, and all cases that were sequenced were <u>negative</u> for mutations in sections of the following candidate genes: FGFR1 exon IIIa 755C->G, FGFR2 (exons IIIa and IIIc,), FGFR3 exon IIIa , and TWIST exon I. Since none of the cases carried mutations, testing for gene-environment interaction was not possible. This study, while small in size, still adds to the growing body of literature examining risk factors for isolated craniosynostosis. It was the first study to sequence isolated craniosynostosis cases for mutations in these genes known to control syndromic forms of craniosynostosis, and because all cases were negative for the mutations, it is highly suggestive that mutations in these genes do not contribute to risk of isolated sagittal craniosynostosis.

Future directions

A larger sample size would have provided more power to detect associations between the environmental risk factors (smoking, alcohol consumption, etc.) and risk to disease, but would be unlikely to add information regarding the mutations studied. Future studies need to examine different mutations, or different regions of these genes, and a larger sample size might help in the assessment of environmental risk factors.

The third research paper, **Methylenetetrahydrofolate Reductase and Congenital Heart Defects: Evidence for Gene-Environment Interaction**, yielded more interesting results. Cases and controls from the Baltimore-Washington Infants Study were genotyped for two MTHFR variants, C677T and A1298C, allowing for traditional case-control analysis.

Findings

There was evidence of gene-environment interaction among the African-American pulmonic stenosis cases with smoking and the MTHFR A1298C marker. There was no evidence of a genotypic effect among the African American PS cases whose mothers were nonsmokers (OR=0.98, 95%CI=0.27-3.42). However, if the mother reported smoking, and the infant carried the C- genotype, there was a significantly increased risk for PS (OR=3.03, 95%CI=1.09-8.52). The attributable fraction (AR), the proportion of cases in the target population attributed to the exposure, was calculated as: AR = p_e (R-1) / [p_e (R-1) + 1]. Estimating the attributable fraction depends on the probability of exposure (p_e). Since in this analysis the combination of genotype and smoking is important, p_e should be the product of the frequency of the exposure (i.e. smoking) and the proportion of susceptible genotypes (i.e. presence one or two MTHFR

A1298C C alleles). The proportion of African-American PS cases that could be avoided by eliminating the exposure (i.e. smoking) was estimated to be 20.1%.

While there were no marginal genotypic effects in any of these four heart defect subgroups, there were several environmental factors that increased risk to congenital heart defects, including: heavy smoking (pulmonic stenosis), heavy alcohol use (pulmonic stenosis, atrial septal defects), and family history of cardiac defects (coarctation of aorta). These results confirmed an earlier report of alcohol consumption increasing risk to ASD, and also underscore that heart defects are heterogeneous in etiology.

Case-only analysis was also performed in an effort to compare the results of the differing types of analyses. The case-only results were very similar to the case-control analyses, and in many instances the odds ratios were the same.

Future directions

Replication studies would need to be performed to determine whether the results from this study represent true associations, or chance findings. Recent studies have shown that folic acid reduces the risk of ventricular septal defects, a group not included in this study of congenital heart defects. This group will be targeted next to test whether MTHFR genotype, environmental factors and/or interactions between them increase risk for this defect.

Birth defects, genetic epidemiology and public health

Birth defects are the leading cause of infant mortality and the fifth leading cause of years of potential life lost in the United States. For 1992, estimated costs of the 17 most clinically important birth defects and cerebral palsy was $8 billion. Costs ranged

from $75,000 to $1.7 million per new case (MMWR, 1995). Congenital anomalies are associated with a considerable amount of economic and social loss. Effective prevention strategies would greatly decrease the economic burden of these birth defects, and avoid substantial social and emotional cost to patients and their families. The following table lists the breakdown of costs for the oral clefts and congenital heart defects.

Table 1: Economic Costs of Oral Clefts, and Heart Defects*

Disease	Total cost	Cost/new case	Incidence/ 10,000 livebirths	Years of Life Lost / 10,000 live births
Oral clefts	697 million	$101,000	1	1.5
Heart defects**	1.3 billion	$1.7 million	11	186

* from MMWR, 1995
** includes truncus arteriosis, single ventricle, transposition of the great arteries, double outlet right ventricle and tetralogy of fallot

Eliminating a modifiable risk factor, such as smoking, would greatly decrease these costs. For example, in this sample of four different types of congenital heart defects, 6.2% of these cases could have been avoided by removing maternal smoking, which would mean a $62,000,000 reduction in costs. Hence, although the attributable fraction for smoking is modest, the savings is substantial. Wyszynski et al.(1997) performed a meta-analysis utilizing the results from 11 studies and found an overall odds ratio of 1.29 (95% CI: 1.18-1.42) for oral clefts in children of women who smoked, and calculated an attributable risk of 11%. This translates into a savings of approximately $77,000,000.

It is evident that birth defects have multiple risk factors, including genetic and environmental causes. Information regarding population distributions of genotypes

associated with birth defects, the implications and feasibility of genetic testing and the efficacy of interventions are largely unknown. With the completion of the Human Genome Project, opportunities for genome screens to locate genes for complex diseases, such as birth defects, have greatly increased as the list of candidate genes has expanded with every advance in our understanding of the biology of normal development. Association studies to examine gene-environment interaction will continue to be an important step in this process. The Centers for Disease Control, recognizing the complex issues surrounding the impact of advances in genetics on disease, created the Office of Genetics and Disease Prevention in 1997, with the goal of monitoring the increasingly important role of genetics in the practice of public health.

The National Institutes of Environmental Health Sciences also acknowledged the importance of interactions between genetics and environment by developing the Environmental Genome Project (EGP; www.niehs.nih.gov/envgenom/concept.htm) . The goal of this project is to identify allelic variants of environmental disease susceptibility genes, develop a database of polymorphisms for these genes, and facilitate population-based studies to further examine gene-environment interaction. The EGP will also attempt to advance techniques for genetic analysis, improve study designs, develop sample repositories and address the ethical, legal and social issues that surround genetic studies.

In 1998, Khoury and Dorman(http://www.cdc.gov/genetics/hugenet/about.htm) emphasized the importance of population-based genetic epidemiologic research. They coined the term HuGE Net, human genetic epidemiology network. The goal of HuGE Net is to address the complexities and implications of the intermingling of molecular

genetics and genetic epidemiology. More specifically, HuGE Net aspires: (1) to establish global collaboration in the development and dissemination of peer_reviewed epidemiologic information on human genes; (2) to develop an updated and accessible knowledge base on the World Wide Web; and (3) to promote the use of this knowledge base by health care providers, researchers, industry, government, and the public for making decisions involving the use of genetic tests and services for disease prevention and health promotion (http://www.cdc.gov/genetics/hugenet/about.htm).

It is important to unravel the genetic and environmental factors that contribute to the etiology of the three birth defects studied in this dissertation. There is an obvious social and economic impact associated with these diseases, and the goal is to find ways to ease this burden. Association studies help identify risk factors (i.e. smoking has now been associated with several birth defects, including oral clefts and heart defects). Repeated replication of the findings may influence public policy makers to implement educational programs for primary prevention and molecular biologists to examine more closely the biological mechanisms by which certain genes increase risk to disease.