Polytomous Logistic Regression as a Tool for Exploring Heterogeneity across Birth Defect Subtypes: An Example Using Anencephaly and Spina Bifida

Philip J. Lupo,[1] Elaine Symanski,[1] D. Kim Waller,[1] Wenyaw Chan,[2] Mark A. Canfield,[3] Peter H. Langlois,[3] and Laura E. Mitchell[1]* and the National Birth Defects Prevention Study

[1]Division of Epidemiology and Disease Control, University of Texas School of Public Health, Houston, Texas
[2]Division of Biostatistics, University of Texas School of Public Health, Houston, Texas
[3]Birth Defects Epidemiology and Surveillance Branch, Texas Department of State Health Services, Austin, Texas

Received 15 January 2010; Revised 5 March 2010; Accepted 15 March 2010

BACKGROUND: In birth defect epidemiology, phenotypic subgroups are often combined into a composite phenotype in an effort to increase statistical power. Although the validity of using composite phenotypes has been questioned, formal evaluations of the underlying assumption of effect homogeneity across component phenotypes have not been conducted. **METHODS:** Polytomous logistic regression was used to assess effect heterogeneity of several generally accepted neural tube defect (NTD) risk factors across the component phenotypes of anencephaly and spina bifida. Data for these analyses were obtained from the National Birth Defects Prevention Study. **RESULTS:** The use of a composite phenotype has the potential to mask associations specific to a component phenotype and in some cases the effect of a variable may be misattributed to the composite phenotype. For example, an association between infant sex and anencephaly (adjusted odds ratio [AOR], 1.5; 95% CI, 1.1–1.9) was masked when data from all NTDs were analyzed (AOR, 1.1; 95% CI, 0.9–1.3), whereas an association with maternal body mass index that was specific to spina bifida (AOR, 1.9; 95% CI, 1.6–2.4) was attributed to all NTDs (AOR, 1.6; 95% CI, 1.4–2.0). Furthermore, conclusions regarding effect heterogeneity based on ad hoc comparisons, rather than some formal assessment, may be vulnerable to considerable subjectivity, as was the case for the association of maternal Hispanic ethnicity with spina bifida (AOR, 1.4; 95% CI, 1.2–1.8) and anencephaly (AOR, 2.0; 95% CI, 1.5–2.8). **CONCLUSIONS:** Polytomous logistic regression provides a useful tool for evaluating putative risk factors for which there is no a priori basis for assuming effect homogeneity across component phenotypes. *Birth Defects Research (Part A) 88:701–705, 2010.* © 2010 Wiley-Liss, Inc.

Key words: neural tube defects; birth defects; heterogeneity; polytomous logistic regression; epidemiology

INTRODUCTION

In birth defect epidemiology, it is common to combine phenotypic subgroups, or component phenotypes, for analysis. Examples of such combined phenotypes, or composite phenotypes, include neural tube defects (NTDs; e.g., anencephaly and spina bifida) and conotruncal heart defects (e.g., tetralogy of Fallot and transposition of the great arteries). In general, component birth defect phenotypes include the same or related structures, and there is evidence of a shared, although not necessarily identical, etiology. For example, anencephaly and spina bifida are neural tube closure defects that co-segregate within families, but affected relatives are more likely to be concordant (e.g., spina bifida/spina bifida) than discordant (e.g., spina bifida/anencephaly) for defect type (Lalouel et al., 1979). Such familial aggregation patterns suggest that the genetic risk factors for spina bifida and anencephaly overlap but are not identical.

This study was supported in part through the cooperative agreement U50/CCU613232 from the Centers for Disease Control and Prevention to the Texas Department of State Health Services Center for Birth Defects Research and Prevention and the Southwest Center for Occupational and Environmental Health Training Grant T42OH008421.
*Correspondence to: Laura E. Mitchell, University of Texas School of Public Health, 1200 Herman Pressler Drive, Houston, TX 77030.
E-mail: Laura.E.Mitchell@uth.tmc.edu
Published online 22 April 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/bdra.20671

Epidemiologic studies also provide evidence that anencephaly and spina bifida share some (Czeizel and Dudas, 1992; Berry et al., 1999; Correa et al., 2008; Mosley et al., 2009) but not all risk factors (Waller et al., 2007; Rasmussen et al., 2008; Stothard et al., 2009). For instance, a woman's risk of having a child with either anencephaly and spina bifida can be significantly reduced by periconceptional, folic acid supplementation (Czeizel and Dudas, 1992). Furthermore, the teratogenic potential of maternal pregestational diabetes has been established for both anencephaly and spina bifida and includes a two- to 10-fold increase in risk among offspring of diabetic women compared to the general population (McLeod and Ray, 2002; Mitchell, 2005). Other factors, such as maternal prepregnancy obesity, which is associated with spina bifida but not anencephaly (Waller et al., 2007; Stothard et al., 2009;), are related to only one of the component phenotypes.

In theory, analyses based on component phenotypes are preferred over those based on the composite phenotype, because the former are likely to be relatively etiologically homogeneous in comparison with the latter (Khoury et al., 1992). However, when the effect of a putative risk factor is similar across the component phenotypes, analyses based on the composite phenotype will provide increased power and effect estimate precision, relative to analyses based on the individual component phenotypes. These benefits are of particular importance in birth defect studies, where sample sizes are limited by the relative rarity of the component phenotypes.

Optimally, analyses of putative birth defect risk factors should be based on individual phenotypes when the effect of the factor is heterogeneous across the component phenotypes and on the composite phenotype when it is not (Khoury et al., 1982). However, because there is often little a priori knowledge regarding potential effect heterogeneity for a given risk factor, separate analyses of the component and composite phenotypes are often conducted, and in general, conclusions regarding effect heterogeneity are based on visual inspection of, and ad hoc comparisons between, effect estimates and confidence intervals (Canfield et al., 2009a; Canfield et al., 2009b; Glynn and Rosner, 2004; Harville et al., 2005; Rowland et al., 2006). Although methods that allow for formal (i.e., statistical) evaluation and modeling of effect heterogeneity across component phenotypes are available (Glynn and Rosner, 2004), they have not been applied in studies of birth defects.

Polytomous logistic regression (PLR), a method commonly used to assess nominal outcomes with more than two levels (Hosmer and Lemeshow, 2000), can also be used to assess effect heterogeneity across component phenotypes and estimate adjusted measures of effect for the component or the composite phenotypes, as appropriate (Glynn and Rosner, 2004). To illustrate the potential utility of this approach in epidemiologic studies of birth defects, we used PLR to formally assess heterogeneity of the effects of several, generally accepted NTD risk factors, across the component phenotypes of anencephaly and spina bifida. As these risk factors have been extensively studied, our purpose was to validate previous conclusions regarding effect heterogeneity, which were based on ad hoc comparisons (Shaw et al., 2003; Canfield et al., 2009a; Canfield et al., 2009b; Stothard et al., 2009) and illustrate the potential utility of PLR for future studies of NTDs and other composite birth defects phenotypes.

METHODS

Study Subjects

The study population included case and control infants/fetuses, from the National Birth Defects Prevention Study (NBDPS), with estimated dates of delivery between October 1, 1997, and December 31, 2004. Details of the NBDPS have been published elsewhere (Yoon et al., 2001). In brief, the NBDPS is an ongoing, population-based, case–control study of major structural birth defects. Cases are or have been ascertained through 10 surveillance systems in the United States (Arkansas, California, Georgia/Centers for Disease Control and Prevention, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah) and include live births, stillbirths, and induced pregnancy terminations. Infants/fetuses with single-gene disorders or chromosome abnormalities are excluded from the NBDPS, and all cases are classified as either isolated or multiple by clinical geneticists (Rasmussen et al., 2003). Control subjects include unaffected live births randomly selected from birth certificates or birth hospitals. The mothers of cases and controls complete a one-hour computer-assisted telephone interview (CATI) that includes sections on maternal conditions and illnesses, lifestyle and behavioral factors, and multivitamin use.

For the present study, cases were limited to include infants/fetuses with a diagnosis of anencephaly (n = 290) or spina bifida (n = 630). NBDPS cases with a diagnosis of both anencephaly and spina bifida (n = 3), an NTD other than anencephaly or spina bifida (craniorachischisis, n = 9; encephalocele, n = 121), and those classified as complex cases (n = 2) were excluded. Cases and controls for whom the mother was reported to have maternal pregestational diabetes (n = 10 cases; n = 29 controls) or use anticonvulsants (n = 15 cases; n = 35 controls) were excluded, because these exposures were relatively rare and are strong risk factors for NTDs.

Data for infant and maternal characteristics that are generally accepted to be associated with NTD risk were obtained from the CATI and included the following: infant sex (male/female), maternal race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and other), maternal use of folic acid supplements in the month prior to conception through the month following conception (any/none), and maternal body mass index (BMI). Maternal BMI (weight [kg] / height [m^2]) was categorized according to the National Heart, Lung and Blood Institute cutoff points as underweight ($<$18.5 kg/m^2), average weight (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2), and obese ($\geq$30.0 kg/m^2).

Statistical Methods

PLR was used to evaluate heterogeneity of the effects of each risk factor across the component NTD phenotypes (Glynn and Rosner, 2004). Two composite phenotypes were considered: (1) NTDs (anencephaly and spina bifida) and (2) isolated NTDs (isolated anencephaly and isolated spina bifida). Each composite phenotype was assessed using a step-up strategy, in which a baseline model with all variables constrained to have equal effect estimates across the component phenotypes was first fitted to the data. Next, a model in which the variable with the highest p value in the baseline model was allowed to vary across

Table 1
Characteristics of Infants with Neural Tube Defects and Control Infants, National Birth Defects Prevention Study, 1997–2004

Characteristic	Spina bifida (%)	Isolated spina bifida (%)	Anencephaly (%)	Isolated anencephaly (%)	Controls (%)
Infant sex					
Male	302 (50.0)	272 (49.4)	101 (40.7)	94 (40.9)	2965 (50.4)
Female	302 (50.0)	279 (50.6)	147 (59.3)	136 (59.1)	2924 (49.6)
Maternal factors:					
BMI (kg/m^2)					
Underweight (<18.5)	22 (3.8)	20 (3.8)	17 (6.4)	16 (6.5)	321 (5.7)
Normal (18.5–24.9)	278 (48.0)	252 (47.9)	145 (54.3)	132 (53.9)	3184 (56.4)
Overweight (25.0–29.9)	134 (23.1)	120 (22.8)	61 (22.9)	56 (22.9)	1249 (22.2)
Obese (≥30)	145 (25.1)	134 (24.5)	44 (16.4)	41 (16.7)	887 (15.7)
Folic acid supplement use[a]					
Nonuser	307 (50.3)	279 (50.1)	137 (48.6)	124 (47.7)	2890 (49.3)
User	304 (49.7)	278 (49.9)	145 (51.4)	136 (52.3)	2978 (50.7)
Race/ethnicity					
Non-Hispanic White	331 (54.2)	304 (54.7)	136 (48.6)	128 (49.6)	3497 (59.6)
Non-Hispanic Black	56 (9.2)	48 (8.6)	24 (8.6)	20 (7.8)	671 (11.4)
Hispanic	190 (31.1)	172 (30.9)	97 (34.6)	88 (34.1)	1321 (22.5)
Other	34 (5.5)	32 (5.8)	23 (8.2)	22 (8.5)	383 (6.5)

[a]Use of folic acid supplements in the month before pregnancy through the first month of pregnancy.

the component phenotypes was fitted to the data and a likelihood ratio test (LRT) comparing the two models was used to evaluate heterogeneity of the effect of the risk factor across the component phenotypes. When the LRT had $p > 0.05$, the effect of the risk factor was considered to be homogeneous, and the variable remained constrained in subsequent models. The variable with the next highest p value was then allowed to vary across the component phenotypes, and the procedure was repeated until a LRT with $p \leq 0.05$ was identified. When $p \leq 0.05$, the effect of the variable was considered to be heterogeneous across the component phenotypes and a model that allowed this variable to vary across the component phenotypes became the baseline for the next set of likelihood ratio tests (Glynn and Rosner, 2004; Wei et al., 2004). A step-down analysis was also performed in which the initial model allowed all variables to vary across the component phenotypes, and LRTs were used to evaluate successive, constrained models. All models included a variable for study center. Analyses were performed using Stata version 10 (StataCorp LP, College Station, TX).

RESULTS

Participation rates for the NBDPS were 71% among mothers with spina bifida–affected pregnancies, 72% among mothers with anencephaly-affected pregnancies, and 68% among control mothers. The distributions of selected demographic characteristics and risk factors for case and control infants/fetuses and mothers are presented in Table 1.

The adjusted odds ratios (AORs) generated from a baseline PLR model in which all effect estimates were allowed to vary across anencephaly and spina bifida (i.e., the initial base model in the step-down approach), and from a baseline model in which all effect estimates were set to be equal for anencephaly and spina bifida (i.e., the initial model in the step-up approach) are presented in Table 2. These effect estimates are analogous to what would be obtained from logistic regression analyses of the component and composite phenotypes, respectively. Visual inspection of the point estimates for anencephaly and spina bifida suggests that the effects of at least a subset of the study variables differ across these component endpoints. For example, the AOR for female sex is 1.5 for anencephaly (95% CI, 1.1–1.9), whereas it is 1.0 for spina bifida (95% CI, 0.9–1.2). However, in some instances ad hoc decisions regarding heterogeneity would be difficult, and formal assessment of effect heterogeneity is not possible based on the data presented in Table 2. This is evident with maternal Hispanic ethnicity, for which the AOR is 1.4 (95% CI, 1.2–1.8) for spina bifida and 2.0 (95% CI, 1.5–2.8) for anencephaly.

Use of the step-up PLR strategy to formally assess the evidence for effect heterogeneity, indicated that, in these data, there was significant heterogeneity in the associations of anencephaly and spina bifida with infant sex (female vs. male: AOR [spina bifida], 1.0; AOR [anencephaly], 1.5, $P_{\text{heterogeneity}} = 0.02$) and maternal BMI (obese vs. normal: AOR [spina bifida], 1.9; AOR [anencephaly], 1.1, $P_{\text{heterogeneity}} = 0.01$), but not with maternal periconceptional use of folic acid supplements, which was not associated with either component endpoint, or maternal race/ethnicity (Table 3). Similar conclusions regarding effect heterogeneity were reached when the step-down procedure was used (Glynn and Rosner, 2004) and when the analyses were restricted to include cases with isolated anencephaly and spina bifida (results not shown).

DISCUSSION

The NBDPS provides a rich data source for assessing birth defect risk factors (Yoon et al., 2001), as well as the opportunity to explore novel methods, such as PLR, that may help to define the complex etiologies of these conditions. However, as with any study, the NBPDS has some limitations that must be considered when evaluating study results. Most relevant to the present study is the potential for differential ascertainment of cases with anencephaly and spina bifida resulting from under-ascer-

Table 2
Polytomous Logistic Regression Models[a] for Spina Bifida, Anencephaly, and the Composite NTD Phenotype, National Birth Defects Prevention Study, 1997–2004

	Component phenotypes		Composite phenotype
Characteristic[b]	Spina bifida (n = 613) AOR (95% CI)	Anencephaly (n = 282) AOR (95% CI)	NTDs (n = 895) AOR (95% CI)
Infant sex			
Male	1.0	1.0	1.0
Female	1.0 (0.9–1.2)	1.5 (1.1–1.9)	1.1 (0.9–1.3)
Maternal factors:			
BMI (kg/m^2)			
Underweight (≤18.5)	0.8 (0.5–1.3)	1.1 (0.6–1.9)	0.9 (0.6–1.3)
Normal (18.5–24.9)	1.0	1.0	1.0
Overweight (25.0–29.9)	1.2 (0.9–1.5)	1.0 (0.7–1.4)	1.2 (0.9–1.4)
Obese (>30)	1.9 (1.6–2.4)	1.1 (0.7–1.5)	1.6 (1.4–2.0)
Folic acid supplement use			
Nonuser[c]	1.0	1.0	1.0
User	1.0 (0.9–1.2)	1.1 (0.8–1.4)	1.1 (0.9–1.2)
Race/ethnicity			
Non-Hispanic White	1.0	1.0	1.0
Non-Hispanic Black	0.8 (0.6–1.1)	0.9 (0.5–1.4)	0.8 (0.6–1.1)
Hispanic	1.4 (1.2–1.8)	2.0 (1.5–2.8)	1.6 (1.3–1.9)
Other	1.0 (0.7–1.5)	1.7 (1.1–2.9)	1.2 (0.9–1.7)

[a]All models included a variable for study center.
[b]All variables were assessed simultaneously.
[c]Use of folic acid supplements in the month before pregnancy through the first month of pregnancy.

tainment of potential subjects among pregnancy terminations. Such under-ascertainment appears likely, given that the participation rates were similar among the anencephaly and spina bifida cases, whereas the number of anencephalic cases (n = 282) was substantially lower than the number of spina bifida cases (n = 613). Previous studies have, however, demonstrated that the potential for this bias is minimal when cases are ascertained through active surveillance (as is the case in the majority of NBDPS centers) and can be further minimized by adjusting for study center (as was done in the current study) to account for differences in case ascertainment (Canfield et al., 2009b). Further, although such potential biases should clearly be considered when interpreting results about potential risk factors, they are less relevant to discussions regarding the selection of statistical methods, such as in the present study.

Because of the rarity of individual birth defects, epidemiologic studies often combine data for presumably related (i.e., component) birth defects in an effort to improve study power. In general, decisions to combine phenotypes are appropriately based on multiple lines of evidence, which may include embryological, developmental, genetic, clinical, and epidemiologic data indicating that the conditions share common, although not necessarily identical, features (Khoury et al., 1982). In acknowledgment of the potential for etiologic heterogeneity across component phenotypes, analyses are often conducted using the composite phenotype and repeated using each of the component phenotypes. However, such analyses do not allow for formal evaluation of effect heterogeneity across the component phenotypes and therefore do not directly address which set or sets of effect measures provide the most valid estimates of effect. Moreover, within a given multivariable model, such analyses require that all effect measures be estimated for either a component or a composite phenotype, when in reality it may be appropriate to estimate some effect measures based on the composite and others based on the component phenotypes.

PLR can be used to assess effect heterogeneity across component phenotypes and, within a given multivariable model, can provide effect measure estimates for either the composite or component phenotypes as appropriate. To illustrate the potential utility of PLR in epidemiologic studies of birth defects, we used this approach to examine several established NTD risk factors. Our results indicate that analyses based on the composite NTD phenotype have the potential to mask associations that are specific to a component phenotype (e.g., infant sex and spina bifida) as well as to misattribute the effects of a variable to the composite phenotype (e.g., BMI and NTDs). Further, these analyses indicate that conclusions regarding effect heterogeneity based on ad hoc comparisons of the effect estimates obtained from each of the component phenotypes may be vulnerable to considerable subjectivity (e.g., maternal race/ethnicity).

Given the considerable amount of research that has been conducted on this set of NTD risk factors, the specific findings of this study are not surprising and are largely consistent with our understanding of NTD epidemiology. (Although the lack of evidence for an association between maternal use of folic acid supplements and both anencephaly and spina bifida may appear to conflict with the majority of studies assessing this relationship, this finding is consistent with a previous NBDPS study [Mosley et al., 2009], which used a slightly different measure of folic acid fortification. It is likely that the lack of an association between folic acid and NTDs in the NBDPS reflects the post–folic acid fortification status of the study population.) However, these analyses serve to illustrate the potential weaknesses of studies that do not formally evaluate and model effect heterogeneity. Moreover, they demonstrate the potential utility of the PLR

Table 3
Final Polytomous Logistic Regression Model[a] Assessing Effect Heterogeneity of Established NTD Risk Factors between Spina Bifida and Anencephaly, National Birth Defects Prevention Study, 1997–2004

Characteristic[b]	Spina bifida (n = 613) AOR (95% CI)	Anencephaly (n = 282) AOR (95% CI)	p value for difference
Infant sex			
Male	1.0	1.0	0.02
Female	1.0 (0.8–1.2)	1.5 (1.1–1.9)	
Maternal factors:			
BMI (kg/m^2)			
Underweight (<18.5)	0.8 (0.5–1.3)	1.1 (0.6–1.9)	0.01
Normal (18.5–24.9)	1.0	1.0	
Overweight (25.0–29.9)	1.2 (1.0–1.5)	1.0 (0.7–1.4)	
Obese (≥30)	1.9 (1.6–2.4)	1.1 (0.7–1.5)	
Folic acid supplement use[c,d]			
Nonuser	1.0		0.92
User, B1-P1	1.1 (0.9–1.2)[e]		
Race/ethnicity[d]			
Non-Hispanic White	1.0		0.16
Non-Hispanic Black	0.8 (0.6–1.1)		
Hispanic	1.6 (1.3–1.9)		
Other	1.2 (0.9–1.7)		

[a]All models included a variable for study center.
[b]All variables were assessed simultaneously.
[c]Use of folic acid supplements in the month before pregnancy through the first month of pregnancy.
[d]Effect estimates are not significantly different across the component phenotypes of spina bifida and anencephaly; therefore the final model provided effect estimates for the composite phenotype (NTDs).
[e]Although these analyses used a different definition of folic acid use than that used by Mosley et al. (2009) in a previous analyses of NBDPS data, the results obtained based on the two different definitions are similar and provide little evidence of an association between maternal, periconceptional use of folic acid and risk of NTDs in offspring in this study population.

approach for analyses aimed at assessing putative disease risk factors for which there may be little if any prior information regarding effect heterogeneity.

ACKNOWLEDGMENTS

We thank the participating centers and families of the National Birth Defects Prevention Study for their valuable contributions. Additionally, we would like to thank Tunu Ramadhani for her invaluable assistance on this project.

REFERENCES

Berry RJ, Li Z, Erickson JD, et al. 1999. Prevention of neural-tube defects with folic acid in China: China-U.S. Collaborative Project for Neural Tube Defect Prevention. N Engl J Med 341(20):1485–1490.
Canfield MA, Marengo L, Ramadhani TA, et al. 2009a. The prevalence and predictors of anencephaly and spina bifida in Texas. Paediatr Perinat Epidemiol 23(1):41–50.
Canfield MA, Ramadhani TA, Shaw GM, et al. 2009b. Anencephaly and spina bifida among Hispanics: maternal, sociodemographic, and acculturation factors in the National Birth Defects Prevention Study. Birth Defects Res A Clin Mol Teratol 85(7):637–646.
Correa A, Gilboa SM, Besser LM, et al. 2008. Diabetes mellitus and birth defects. Am J Obstet Gynecol 199(3):237.e1–239.e9.
Czeizel AE, Dudas I. 1992. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. N Engl J Med 327(26):1832–1835.
Glynn RJ, Rosner B. 2004. Methods to evaluate risks for composite end points and their individual components. J Clin Epidemiol 57(2):113–122.
Harville EW, Wilcox AJ, Lie RT, et al. 2005. Cleft lip and palate versus cleft lip only: are they distinct defects?. Am J Epidemiol 162(5):448–453.
Hosmer DW, Lemeshow S. 2000. Applied logistic regression. 2nd ed. Hoboken, NJ: John Wiley & Sons.
Khoury MJ, Erickson JD, James LM, et al. 1982. Etiologic heterogeneity of neural tube defects. II. Clues from family studies. Am J Hum Genet 34(6):980–987.
Khoury MJ, Moore CA, James LM, et al. 1992. The interaction between dysmorphology and epidemiology: methodologic issues of lumping and splitting. Teratology 45(2):133–138.
Lalouel JM, Morton NE, Jackson J, et al. 1979. Neural tube malformations: complex segregation analysis and calculation of recurrence risks. J Med Genet 16(1):8–13.
McLeod L, Ray JG. 2002. Prevention and detection of diabetic embryopathy. Community Genet 5(1):33–39.
Mitchell LE. 2005. Epidemiology of neural tube defects. Am J Med Genet C Semin Med Genet 135(1):88–94.
Mosley BS, Cleves MA, Siega-Riz AM, et al. 2009. Neural tube defects and maternal folate intake among pregnancies conceived after folic acid fortification in the United States. Am J Epidemiol 169(1):9–17.
Rasmussen SA, Chu SY, Kim SY, et al. 2008. Maternal obesity and risk of neural tube defects: a metaanalysis. Am J Obstet Gynecol 198(6):611–619.
Rasmussen SA, Olney RS, Holmes LB, et al. 2003. Guidelines for case classification for the National Birth Defects Prevention Study. Birth Defects Res A Clin Mol Teratol 67(3):193–201.
Rowland CA, Correa A, Cragan JD, et al. 2006. Are encephaloceles neural tube defects? Pediatrics 118(3):916–923.
Shaw GM, Carmichael SL, Kaidarova Z, et al. 2003. Differential risks to males and females for congenital malformations among 2.5 million California births, 1989–1997. Birth Defects Res A Clin Mol Teratol 67(12):953–958.
Stothard KJ, Tennant PWG, Bell R, et al. 2009. Maternal overweight and obesity and the risk of congenital anomalies. JAMA 301(6):636–650.
Waller DK, Shaw GM, Rasmussen SA, et al. 2007. Prepregnancy obesity as a risk factor for structural birth defects. Arch Pediatr Adolesc Med 161(8):745–750.
Wei EK, Giovannucci E, Wu K, et al. 2004. Comparison of risk factors for colon and rectal cancer. Int J Cancer 108(3):433–442.
Yoon PW, Rasmussen SA, Lynberg MC, et al. 2001. The National Birth Defects Prevention Study. Public Health Rep 116(Suppl 1):32–40.