HHS Public Access

Author manuscript

Birth Defects Res A Clin Mol Teratol. Author manuscript; available in PMC 2015 September 14.

Published in final edited form as:

Birth Defects Res A Clin Mol Teratol. 2012 March ; 94(3): 141–146. doi:10.1002/bdra.22883.

Working Towards a Risk Prediction Model for Neural Tube Defects

A.J. Agopian[1], **Philip J. Lupo**[1], **Sarah C. Tinker**[2], **Mark A. Canfield**[3], **Laura E. Mitchell**[1,*], and **the National Birth Defects Prevention Study**

[1]Human Genetics Center, Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, Texas

[2]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, Georgia

[3]Birth Defects Epidemiology and Surveillance Branch, Texas Department of State Health Services, Austin, Texas

Abstract

BACKGROUND—Several risk factors have been consistently associated with neural tube defects (NTDs). However, the predictive ability of these risk factors in combination has not been evaluated.

METHODS—To assess the predictive ability of established risk factors for NTDs, we built predictive models using data from the National Birth Defects Prevention Study, which is a large, population-based study of nonsyndromic birth defects. Cases with spina bifida or anencephaly, or both (n = 1239), and controls (n = 8494) were randomly divided into separate training (75% of cases and controls) and validation (remaining 25%) samples. Multivariable logistic regression models were constructed with the training samples. The predictive ability of these models was evaluated in the validation samples by assessing the area under the receiver operator characteristic curves. An ordinal predictive risk index was also constructed and evaluated. In addition, the ability of classification and regression tree (CART) analysis to identify subgroups of women at increased risk for NTDs in offspring was evaluated.

RESULTS—The predictive ability of the multivariable models was poor (area under the receiver operating curve: 0.55 for spina bifida only, 0.59 for anencephaly only, and 0.56 for anencephaly and spina bifida combined). The predictive abilities of the ordinal risk indexes and CART models were also low.

CONCLUSION—Current established risk factors for NTDs are insufficient for population-level prediction of a women's risk for having affected offspring. Identification of genetic risk factors and novel nongenetic risk factors will be critical to establishing models, with good predictive ability, for NTDs.

*Correspondence to: Laura E. Mitchell, University of Texas School of Public Health, 1200 Herman Pressler Dr., Houston, TX 77030.

The findings and conclusions in this report are those of the authors and do not necessarily reflect those of the Centers for Disease Control and Prevention.

Keywords

spina bifida; neural tube; anencephaly; congenital abnormalities; prediction; risk score

INTRODUCTION

Spina bifida and anencephaly are severe neural tube defects (NTDs) that result from abnormalities in the development of the spinal cord and brain, respectively. These conditions are among the more common birth defects, with an estimated 6 in 10,000 live births in the United States affected by spina bifida or anencephaly (Parker et al., 2010). The exact pathogenesis of these conditions is unknown. However, several risk factors for spina bifida and anencephaly have been consistently identified, including: family history (reviewed in Mitchell, 2005; Lupo et al., 2010a), female sex (Seller, 1987; Shaw et al., 2003), as well as maternal folate status and folic acid intake (Bailey et al., 2003; Toriello, 2011), Hispanic ethnicity (Canfield et al., 2006; Feuchtbaum et al., 1999), obesity (Waller et al., 1994; Shaw et al., 1996), pregestational and gestational diabetes (Becerra et al., 1990; Correa et al., 2008; Canfield et al., 2009; Lupo et al., 2010a), anticonvulsant use (Lammer et al., 1987; Rosa, 1991), and hyperthermia (Milunsky et al., 1992; Shaw et al., 1998; Duong et al., 2011). While these associations are established or strongly suspected, the predictive ability of these risk factors has not been evaluated in combination.

Because spina bifida and anencephaly share some epidemiologic characteristics, result from similar embryological processes (i.e., failure of the neural tube to close properly), and co-aggregate within families, many studies of NTDs have combined the two component phenotypes (i.e., spina bifida, anencephaly) into a single, composite case definition (composite NTDs; Mitchell, 2005). However, based on previous evidence that NTDs share some but not all risk factors (Feuchtbaum et al., 1999; Rowland et al., 2006; Waller et al., 2007; Correa et al., 2008; Rasmussen et al., 2008; Lupo et al., 2010b), it is prudent to evaluate risk using both the component and composite NTD case definitions.

Whereas studies that focus on assessing associations between individual exposures and birth defects provide insights into etiology, studies that assess predictive ability of risk factors can help to identify high-risk groups in the population (e.g., to target specific intervention strategies to a particular subgroup). Predictive modeling has been used in epidemiologic studies of other conditions (e.g., cardiovascular disease, stroke, cancer; D'Agostino et al., 1994, 2008; Trepanier et al., 2004; Spitz et al., 2007), but it has not been used widely for the purpose of predicting risk to women for birth defects in offspring. Therefore, in the present study we evaluated the ability of various combinations of established or strongly suspected risk factors for NTDs to discriminate between women with offspring with spina bifida or anencephaly and controls in the National Birth Defects Prevention Study, assessing both component and composite NTD case definitions.

MATERIALS AND METHODS

Study Subjects

The present study is based on data from the National Birth Defects Prevention Study (NBDPS), an ongoing, population-based case-control study of birth defects, including NTDs. NBDPS data were collected from 10 birth defects surveillance systems: Arkansas (AR), California (CA), Georgia (GA), Iowa (IA), Massachusetts (MA), New Jersey (NJ), New York (NY), North Carolina (NC), Texas (TX), and Utah (UT). Approval from the institutional review board for each study site was obtained.

Details of the methods for NBDPS subject recruitment and data collection have been published elsewhere (Yoon et al., 2001). Cases were ascertained through population-based birth defects surveillance systems. Cases included live births (all sites), fetal deaths ≥20 weeks (AR, CA, GA, IA, NC, NY, MA, TX, and UT only), and elective pregnancy terminations (AR, CA, GA, IA, NC, NY, TX, and UT only). Medical records for case infants were abstracted, and NBDPS clinical geneticists reviewed these data to confirm diagnoses of eligible defects (Rasmussen et al., 2003). Cases with single-gene disorders or chromosome abnormalities (i.e., syndromic cases) were excluded. Controls were randomly selected from live born infants without major birth defects delivered to women in the study regions. Controls were identified through birth certificate data or hospital birth logs. Participating mothers of cases (live births, fetal deaths, and elective pregnancy terminations) and controls completed a computer-assisted telephone interview in which they were asked about exposures before and during their pregnancy, pregnancy and family histories, maternal conditions, and lifestyle and behavioral factors.

We included data for NBDPS cases with spina bifida or anencephaly and controls with estimated dates of delivery between October 1, 1997, and December 31, 2007, in our analyses. To limit heterogeneity within the case groups, we only included cases with isolated spina bifida or anencephaly. NBDPS clinical geneticists classified cases as having isolated spina bifida or anencephaly if no additional malformations were present, or if only malformations that were likely to be secondary to spina bifida or anencephaly (e.g., talipes equinovarus, hip dislocation, hydrocephalus) were present (Rasmussen et al., 2003).

Predictor Variables

Only variables that are established or strongly suspected risk factors for NTDs were considered in this analysis (reviewed in Lupo et al., 2010a). To maximize the chance to build models with good predictive ability, we considered a broad list of risk factors, some of which are stronger or more established risk factors than others. Potential predictor variables were obtained from the interview data; these included the following dichotomous (yes or no) variables: maternal Hispanic ethnicity, obesity before pregnancy (body mass index ≥30.0), lack of any folic acid supplementation (folic acid, multivitamin, or prenatal supplement), low dietary folate intake (defined in the next paragraph), anticonvulsant medication use, type I or II diabetes diagnosed before index pregnancy, gestational diabetes during index pregnancy, any hot tub or sauna use, and family history of NTDs in a first or second-degree

relative of the case. In addition, infant sex was included to assess differences in risk between mothers of males and females.

Folic acid supplementation, anticonvulsant use, and hot tub or sauna use were defined based on use during the month before pregnancy and the first month of pregnancy (B1–P1). Dietary folate intake during the year before pregnancy was assessed based on interview questions from a shortened version of the Willett food frequency questionnaire (Willett and Lenart, 1998). Total dietary folate intake was expressed as dietary folate equivalents, which account for the increased bioavailability of folate from foods fortified with folic acid (Institute of Medicine, 1999). Low dietary folate intake was defined by a dietary folate equivalent value in the lowest quartile, based on the distribution of dietary folate equivalent among controls (i.e., ≤333.6 μg).

In post hoc analyses, alternative, categorical definitions of race or ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, other), body mass index (<18.5 kg/m^2, 18.5–24.9 kg/m^2, 25.0–29.9 kg/m^2, ≥30 kg/m^2), and dietary folate intake (based on all four quartiles defined by dietary folate equivalents among controls; i.e., ≤333.6 μg, 333.7–495.9 μg, 496.0–720.5 μg, >720.5 μg) were also considered. All analyses were also repeated using data only from the NBDPS surveillance sites that ascertain elective pregnancy terminations (AR, CA, GA, IA, and TX only).

Statistical Methods

Frequency distributions of maternal demographic and behavioral factors were tabulated for case and control infants. Predictive statistical models were built and evaluated separately considering three case definitions: a composite NTD definition, including cases with anencephaly or spina bifida, and two component case definitions, one including only cases with spina bifida and one including only cases with anencephaly. For each of these case definitions, cases and controls were randomly divided into separate training (constituting 75% of cases and controls) and validation (constituting the remaining 25% of cases and controls) samples. (These proportions are based on methods described in previous studies [Barlow et al., 2006; Spitz et al., 2007].) Some cases may have been represented in the composite NTD training sample and the component validation sample or vice versa. The training samples were used to guide model development, and the validation samples were used to evaluate the predictive ability of the models. All analyses were conducted using SAS (version 9.2; SAS, Inc., Cary, NC), unless noted otherwise.

Multivariable Models

Initially, multivariable models included main effects terms for all of the potential predictor variables, and a manual backwards stepwise selection procedure was used to select final models. Variables with nonsignificant p values (i.e., p _0.05) were consecutively removed by descending p value until only variables with significant p values (i.e. $p < 0.05$) remained in the final models. The final models were then evaluated in the validation samples. Hosmer-Lemeshow goodness-of-fit tests were conducted for each model to assess model fit in the validation sample. This test assesses model calibration by evaluating the correspondence between observed and predicted probability rates for groups of cases

(Royston et al., 2009; Farooq et al., 2011). In addition, receiver operating characteristic (ROC) curves were constructed by plotting the sensitivity and specificity of each model in the validation sample. To determine each model's ability to discriminate between cases and controls, the area under the curve (AUC) statistic and 95% confidence intervals were calculated for each ROC curve. An AUC statistic of 0.5 indicates chance prediction (e.g., comparable to a coin toss), and a value of 0.7 or higher indicates good predictive ability (Bewick et al., 2004).

Ordinal Risk Index

To assess predictive ability of a simple index with potential clinical utility, we generated an ordinal risk index. This index was based only on variables that were included in the phenotype-specific regression models and was determined for each case and control mother by increasing a score of zero by one unit for each risk factor present for that mother. For example, the offspring of an obese, Hispanic mother with a family history of NTDs and no other risk factors would have a risk score of 3. The predictive ability of a univariate logistic regression model that included this ordinal risk index variable was assessed in the validation samples by evaluating Hosmer-Lemeshow goodness-of-fit tests and AUC statistics.

Classification and Regression Tree Analysis

To identify subgroups of women at increased risk for having offspring with an NTD, classification and regression tree (CART) analysis (Breiman, 1984) was used. CART is a nonparametric statistical procedure that can identify mutually exclusive population subgroups with common characteristics that are associated with the likelihood of an outcome (Lemon et al., 2003). This process is accomplished by partitioning data into maximally homogeneous subgroups to generate decision tree classification algorithms (Breiman, 1984). Unlike typical regression-based predictive modeling approaches, which are often used to estimate the average effect of independent variables on a dependent variable over the entire population, CART analysis can examine differing effects within population subgroups (Katz, 2006). Thus, CART is well-suited to evaluate and interpret complex interactions involving several variables (Lemon et al., 2003).

CART analyses were conducted using data from the complete datasets (i.e., training and validation sets combined). CART models were created using the Salford Predictive Model Builder program (version 6.6, Salford Systems, San Diego, CA), using the Gini impurity function to define decision trees. All the potential predictor variables were evaluated for possible inclusion in the CART models. A 10-fold cross validation was used to obtain reliable estimates of classification accuracy (Zhang, 1998) and maximal tree depth was set to four levels. The predictive ability of the CART models was assessed by evaluating AUC statistics, which were generated by the Salford Predictive Model Builder program.

RESULTS

For the period October 1, 1997, through December 31, 2007, the NBDPS included 1239 cases with isolated spina bifida or anencephaly (spina bifida, n=836, 67.5%; anencephaly,

n=403, 32.5%) and 8494 controls. The distribution of NTD risk factors by case status and NTD phenotypes are presented in Table 1.

The final multivariable predictive logistic regression model for the composite phenotype included: family history of NTDs, maternal Hispanic ethnicity, obesity, anticonvulsant use, and lower dietary folate intake (Table 2). The final model for spina bifida also included these variables. Predictors for anencephaly were female infant sex, family history of NTDs, and maternal Hispanic ethnicity and lower dietary folate intake. For the three factors predictive of both component phenotypes (i.e., family history of NTDs, maternal Hispanic ethnicity, and lower dietary folate intake), the magnitude of associations were stronger for anencephaly than for spina bifida.

Validation statistics for the final multivariable models are summarized in Table 3. Nonsignificant Hosmer-Lemeshow goodness-of-fit tests indicated that all three of these multivariable models were well calibrated throughout the entire range of probabilities, suggesting that the models fit the data well. The multivariable model AUC scores for composite NTDs, spina bifida, and anencephaly were 0.56, 0.55, and 0.59, respectively. The AUC scores did not increase when categorical variables (rather than dichotomous variables) for maternal race or ethnicity, obesity, or dietary folate intake were considered in these multivariable models (data not shown).

The mean values of the ordinal risk index score for cases with composite NTDs, spina bifida, and anencephaly were similar: 0.87 (SD, 0.79), 0.87 (SD, 0.80), and 0.89 (SD, 0.77), respectively. The mean value for controls was lower, 0.63 (SD, 0.69; data not shown). Results from the three univariable logistic regression models with the ordinal risk index variable are presented in Table 2. Effect sizes for the association between the ordinal risk index score and each outcome were similar between these models, with the strongest effect size present for anencephaly (odds ratio, 1.62; 95% confidence interval, 1.41–1.85). For these models, Hosmer-Lemeshow goodness-of-fit tests indicated that the models fit the data well (i.e., model calibration was not poor). However, AUC scores were 0.56, 0.55, and 0.57 for composite NTDs, spina bifida, and anencephaly, respectively (Table 3), indicating that the models provided poor discrimination between case and control mothers.

CART analyses were also conducted separately for each of the three case groups, and a decision tree was generated for each case group (data not shown). For composite NTDs and spina bifida, the variables represented in these decision trees were similar to those in the corresponding final multivariable logistic regression model. However, lower dietary folate intake was not present in either decision tree, and anticonvulsant use was not present in the tree for composite NTDs. The decision tree for anencephaly included only NTD family history and maternal Hispanic ethnicity. The AUC scores for the decision trees for composite NTDs, spina bifida, and anencephaly, respectively, were 0.58, 0.58, and 0.57 (data not shown). All analyses were also repeated using data only from the NBDPS surveillance sites that ascertain elective pregnancy terminations, and the predictive ability of these models were similar to those in the main analytic group.

DISCUSSION

Based on our analyses of these NBDPS data, it appears that the known and strongly suspected risk factors for NTDs poorly discriminate between mothers of cases and controls, as indicated by AUC scores less than 0.6. An AUC value of 0.5 corresponds to chance prediction, comparable to a coin toss, whereas a value of 0.7 or higher indicates good predictive ability (Bewick et al., 2004).

Knowledge of NTD risk factors can be helpful for individual risk counseling. For example, an individual with a history of a previous NTD-affected pregnancy would have a substantially increased risk for NTDs in future pregnancies. However, in the present analyses, we demonstrate that known NTD risk factors are not sufficient for population-based screening to identify women at high-risk for spina bifida or anencephaly in offspring. This finding may be due in part to the rarity of strong NTD risk factors (e.g., NTD family history, anticonvulsant use) and the fact that the more common risk factors (e.g., infant sex) tend to be weakly associated with risk. Given the rarity of NTDs, the likelihood of developing models with good predictive ability may be low in the absence of relatively common risk factors with strong effects. Regardless, current recommendations (e.g., changing epilepsy therapy from high-risk to low-risk drugs) are warranted, although the effects on population risk will likely be negligible.

The protective effects of folic acid supplementation or higher dietary folate intake, or both, have not been as obvious in recent studies as those seen before mandatory fortification of food products with folic acid in the United States in 1998 (e.g., fortification may mask the effects of supplementation; Mosley et al., 2009; Ahrens et al., 2011). Thus, the predictive ability of the lack of folic acid supplementation and low dietary folate intake variables in the present study may have been lower than it would have been among cases born exclusively before 1998.

The low predictive ability of known and suspected NTD risk factors suggests that postfortification factors responsible for the majority of NTD risk remain unidentified. Therefore, these results highlight the importance of pursuing research strategies to identify genetic risk factors as well as novel nongenetic risk factors for NTDs. However, to improve population predictive models, novel factors may need to be relatively common and have relatively strong effects, and it is possible that such factors do not exist.

There are a few limitations of our analyses. The response rate for the study was 65.8% among controls and 67.6% among cases with spina bifida or anencephaly (data not shown). Because data on nonresponders were not available, we could not assess the possibility of bias related to response rate. Recall bias is a common concern in retrospective studies; however, it is unlikely that recall bias could explain the observed poor predictive ability of the models, because overreporting of exposures in cases compared with controls would be expected to increase the strength of associations, which would increase the predictive ability of the models. Despite the large overall sample size, the sample size for cases with anencephaly was relatively small, and this may have limited our ability to develop and evaluate models within this subgroup. It is also possible that an expanded list of potential

predictors, including genetic factors, would have enhanced the predictive ability of models. However, major genes involved in NTD risk have not yet been established (Lupo et al., 2010a).

This study also has several strengths. The NBDPS is one of the largest studies of NTDs, with data from diverse populations across the United States ascertained from population-based surveillance systems. Detailed clinical information was abstracted from medical records and reviewed by clinical geneticists, to ensure accurate diagnoses and exclude cases with known single gene disorders and chromosome abnormalities. To limit potential etiologic heterogeneity within case definitions, cases with multiple birth defects were also excluded and case definitions were considered that evaluated cases with spina bifida and anencephaly separately and together.

Comprehensive predictive models and risk scores have been developed and used for a variety of other chronic conditions (e.g. cardiovascular disease, stroke, several types of cancer, and many other common conditions; D'Agostino et al., 1994, 2008; Trepanier et al., 2004; Spitz et al., 2007), but the use of predictive risk models based on established risk factors have not been widely used for the purpose of predicting risk for NTDs or other birth defects. These results serve as a first step towards developing comprehensive predictive models for NTDs with good predictive ability. Our findings indicate that research focusing on identifying novel genetic and nongenetic risk factors for spina bifida and anencephaly is needed. The availability of prediction models with good predictive ability could lead to future birth defects prevention, by providing screening tools for individuals at high-risk for NTDs in offspring, as well as by guiding development of intervention strategies specific to high-risk subgroups of women with single or multiple risk factors.

Acknowledgments

Supported by the Texas Center for Birth Defects Research and Prevention, under a cooperative agreement (#5U01DD000494-03) from the Centers for Disease Control and Prevention with the Texas Department of State Health Services

References

Ahrens K, Yazdy MM, Mitchell AA, et al. Folic acid intake and spina bifida in the era of dietary folic acid fortification. Epidemiology. 2011; 22:731–737. [PubMed: 21659881]

Bailey LB, Rampersaud GC, Kauwell GP, et al. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: evolving science. J Nutr. 2003; 133:1961S–1968S. [PubMed: 12771346]

Barlow WE, White E, Ballard-Barbash R, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006; 98:1204–1214. [PubMed: 16954473]

Becerra JE, Khoury MJ, Cordero JF, et al. Diabetes mellitus during pregnancy and the risks for specific birth defects: a population-based case-control study. Pediatrics. 1990; 85:1–9. [PubMed: 2404255]

Bewick V, Cheek L, Ball J, et al. Statistics review 13: receiver operating characteristic curves. Crit Care. 2004; 8:508–512. [PubMed: 15566624]

Breiman, L. Classification and regression trees. Belmont, Calif: Wadsworth International Group; 1984.

Canfield MA, Honein MA, Yuskiv N, et al. National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999–2001. Birth Defects Res A Clin Mol Teratol. 2006; 76:747–756. [PubMed: 17051527]

Canfield MA, Ramadhani TA, Shaw GM, et al. 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. 2009; 85:637–646. [PubMed: 19334286]

Correa A, Gilboa SM, Besser LM, et al. Diabetes mellitus and birth defects. Am J Obstet Gynecol 199. 2008; 237:e231–e239.

D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008; 117:743–753. [PubMed: 18212285]

D'Agostino RB, Wolf PA, Belanger AJ, et al. Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. Stroke. 1994; 25:40–43. [PubMed: 8266381]

Duong HT, Shahrukh Hashmi S, Ramadhani T, et al. Maternal use of hot tub and major structural birth defects. Birth Defects Res A Clin Mol Teratol. 2011; 91:836–841. [PubMed: 21648056]

Farooq V, Brugaletta S, Vranckx P, et al. A guide to interpreting and assessing the performance of prediction models. Euro Intervention. 2011; 6:909–912. [PubMed: 21330235]

Feuchtbaum LB, Currier RJ, Riggle S, et al. Neural tube defect prevalence in California (1990–1994): eliciting patterns by type of defect and maternal race/ethnicity. Genet Test. 1999; 3:265–272. [PubMed: 10495925]

Institute of Medicine. Dietary reference intakes: thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, pantothenic acid, biotin, and choline. Washington, DC: National Academy Press; 1999.

Katz, MH. Multivariable analysis: a practical guide for clinicians. 2. Cambridge, New York: Cambridge University Press; 2006.

Lammer EJ, Sever LE, Oakley GP Jr. Teratogen update: valproic acid. Teratology. 1987; 35:465–473. [PubMed: 3114906]

Lemon SC, Roy J, Clark MA, et al. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med. 2003; 26:172–181. [PubMed: 14644693]

Lupo, PJ.; Etheredge, AJ.; Agopian, A.; Mitchell, LE. Epidemiology of neural tube defects. In: Sheiner, E., editor. Perinatal epidemiology. Hauppauge, NY: Nova Science Publishing; 2010a. p. 411-438.

Lupo PJ, Symanski E, Waller DK, et al. Polytomous logistic regression as a tool for exploring heterogeneity across birth defect subtypes: an example using anencephaly and spina bifida. Birth Defects Res A Clin Mol Teratol. 2010b; 88:701–705. [PubMed: 20740595]

Milunsky A, Ulcickas M, Rothman KJ, et al. Maternal heat exposure and neural tube defects. JAMA. 1992; 268:882–885. [PubMed: 1640616]

Mitchell LE. Epidemiology of neural tube defects. Am J Med Genet C Semin Med Genet. 2005; 135C: 88–94. [PubMed: 15800877]

Mosley BS, Cleves MA, Siega-Riz AM, et al. Neural tube defects and maternal folate intake among pregnancies conceived after folic acid fortification in the United States. Am J Epidemiol. 2009; 169:9–17. [PubMed: 18953063]

Parker SE, Mai CT, Canfield MA, et al. Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004–2006. Birth Defects Res A Clin Mol Teratol. 2010; 88:1008–1016. [PubMed: 20878909]

Rasmussen SA, Chu SY, Kim SY, et al. Maternal obesity and risk of neural tube defects: a metaanalysis. Am J Obstet Gynecol. 2008; 198:611–619. [PubMed: 18538144]

Rasmussen SA, Olney RS, Holmes LB, et al. Guidelines for case classification for the National Birth Defects Prevention Study. Birth Defects Res A Clin Mol Teratol. 2003; 67:193–201. [PubMed: 12797461]

Rosa FW. Spina bifida in infants of women treated with carbamazepine during pregnancy. N Engl J Med. 1991; 324:674–677. [PubMed: 1994251]

Rowland CA, Correa A, Cragan JD, et al. Are encephaloceles neural tube defects? Pediatrics. 2006; 118:916–923. [PubMed: 16950981]

Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009; 338:b604. [PubMed: 19336487]

Seller MJ. Neural tube defects and sex ratios. Am J Med Genet. 1987; 26:699–707. [PubMed: 3551611]

Shaw GM, Carmichael SL, Kaidarova Z, et al. Differential risks to males and females for congenital malformations among 2.5 million California births, 1989–1997. Birth Defects Res A Clin Mol Teratol. 2003; 67:953–958. [PubMed: 14745913]

Shaw GM, Todoroff K, Velie EM, et al. Maternal illness, including fever and medication use as risk factors for neural tube defects. Teratology. 1998; 57:1–7. [PubMed: 9516745]

Shaw GM, Velie EM, Schaffer D, et al. Risk of neural tube defectaffected pregnancies among obese women. JAMA. 1996; 275:1093–1096. [PubMed: 8601928]

Spitz MR, Hong WK, Amos CI, et al. A risk model for prediction of lung cancer. J Natl Cancer Inst. 2007; 99:715–726. [PubMed: 17470739]

Toriello HV. Policy statement on folic acid and neural tube defects. Genet Med. 2011; 13:593–596. [PubMed: 21552133]

Trepanier A, Ahrens M, McKinnon W, et al. Genetic cancer risk assessment and counseling: recommendations of the national society of genetic counselors. J Genet Couns. 2004; 13:83–114. [PubMed: 15604628]

Waller DK, Mills JL, Simpson JL, et al. Are obese women at higher risk for producing malformed offspring? Am J Obstet Gynecol. 1994; 170:541–548. [PubMed: 8116710]

Waller DK, Shaw GM, Rasmussen SA, et al. Prepregnancy obesity as a risk factor for structural birth defects. Arch Pediatr Adolesc Med. 2007; 161:745–750. [PubMed: 17679655]

Willett, W.; Lenart, E. Nutritional epidemiology. New York: Oxford University Press; 1998. Reproducibility and validity of food-frequency questionnaires; p. 101-147.

Yoon PW, Rasmussen SA, Lynberg MC, et al. The National Birth Defects Prevention Study. Public Health Rep. 2001; 116(Suppl 1):32–40. [PubMed: 11889273]

Zhang H. Classification trees for multiple binary responses. J Am Stat Assoc. 1998; 93:180–193.

Table 1

Infant and maternal characteristics of spina bifida and anencephaly cases and controls, National Birth Defects Prevention Study, 1997–2007

	Spina Bifida[a] (N=836)			Anencephaly[a] (N=403)			Controls[a] (N=8,494)	
Exposure	**N**	**(%)**	***P*-value[b]**	**N**	**(%)**	***P*-value[b]**	**N**	**(%)**
Infant factors:								
Sex								
Male	414	(50.1)	0.69	163	(45.2)	0.03	4,314	(50.8)
Female	412	(49.9)		198	(54.8)		4,172	(49.2)
Family history of NTDs[c]								
Yes	19	(2.3)	<0.001	10	(2.5)	<0.001	32	(0.4)
No	817	(97.7)		393	(97.5)		8,462	(99.6)
Maternal factors:								
Race/ethnicity								
Non-Hispanic White	451	(54.1)	<0.001	194	(48.4)	<0.001	4,958	(58.6)
Non-Hispanic Black	65	(7.8)		28	(7.0)		935	(11.0)
Hispanic	265	(31.8)		144	(35.9)		1,918	(22.7)
Other	53	(6.4)		35	(8.7)		653	(7.7)
BMI (kg/m^2)								
Underweight (<18.5)	28	(3.6)	<0.001	24	(6.4)	0.30	438	(5.4)
Normal (18.5–24.9)	356	(46.0)		191	(50.8)		4,474	(55.1)
Overweight (25.0–29.9)	185	(23.9)		87	(23.1)		1,854	(22.8)
Obese (≥30)	205	(26.5)		74	(19.7)		1,360	(16.7)
Folic acid supplementation[d,e]								
Yes	409	(49.1)	0.36	208	(51.6)	0.73	4,293	(50.7)
No	424	(50.9)		195	(48.4)		4,167	(49.3)
Low dietary folate intake (daily µg)								
Quartile 4 (>720.5)	193	(23.3)	0.06	92	(22.9)	<0.001	2,107	(25.0)
Quartile 3 (496.0–720.5)	201	(24.3)		75	(18.7)		2,106	(25.0)
Quartile 2 (333.7–495.9)	202	(24.4)		103	(25.6)		2,107	(25.0)
Quartile 1(≤333.6)	232	(28.0)		132	(32.8)		2,107	(25.0)

	Spina Bifida[a] (N=836)			Anencephaly[a] (N=403)			Controls[a] (N=8,494)	
Exposure	**N**	**(%)**	***P*-value[b]**	**N**	**(%)**	***P*-value[b]**	**N**	**(%)**
Anticonvulsant use[e]								
Yes	16	(1.9)	<0.001	6	(1.5)	0.14	65	(0.8)
No	816	(98.1)		393	(98.5)		8,369	(99.2)
Diabetes before index pregnancy								
Yes	3	(0.4)	0.63	8	(2.0)	0.005	51	(0.6)
No	831	(99.6)		395	(98.0)		8,423	(99.4)
Gestational diabetes during index pregnancy								
Yes	43	(5.2)	0.13	13	(3.2)	0.41	343	(4.0)
No	791	(94.8)		390	(96.8)		8,131	(96.0)
Hot tub or sauna use[e]								
Yes	40	(4.9)	0.25	24	(6.1)	0.84	488	(5.9)
No	780	(95.1)		370	(93.9)		7,852	(94.1)

[a] Characteristic totals may not equal case group totals due to missing data

[b] *P*-value from chi-square test (or Fisher exact test when an expected cell count was <5) of the comparison with controls

[c] In first or second-degree relative

[d] Any use of folic acid, multivitamin, or prenatal vitamin supplementation

[e] During the month before pregnancy and the first month of pregnancy (B1-P1)

Table 2

Adjusted odds ratios in complete sample (training and validation samples combined) for multivariable logistic regression and ordinal risk index models, National Birth Defects Prevention Study, 1997–2007

Variable	Adjusted odds ratio (95% CI)		
	Spina bifida	**Anencephaly**	**Anencephaly + spina bifida**
Multivariable models:			
Family history of NTDs[a]	6.93 (3.84–12.50)	8.36 (4.04–17.30)	7.17 (4.24–12.10)
Hispanic ethnicity	1.54 (1.30–1.82)	2.17 (1.73–2.71)	1.67 (1.45–1.93)
Obesity	1.77 (1.49–2.11)	-	1.56 (1.35–1.81)
Anticonvulsant use[b]	2.37 (1.31–4.28)	-	2.29 (1.37–3.82)
Low dietary folate intake[c]	1.22 (1.03–1.44)	1.57 (1.25–1.98)	1.33 (1.16–1.53)
Female infant sex	-	1.26 (1.02–1.56)	-
Ordinal risk index model:			
Ordinal risk index score[d]	1.56 (1.41–1.72)	1.62 (1.41–1.85)	1.57 (1.45–1.71)

[a] In first or second-degree relative

[b] During the month before pregnancy and the first month of pregnancy (B1-P1)

[c] Based on the lowest quartile of dietary folate equivalent level in controls

[d] Unadjusted odds ratio for each one unit increase in ordinal risk index

Birth Defects Res A Clin Mol Teratol. Author manuscript; available in PMC 2015 September 14.

Table 3

Model validation statistics in validation sample for multivariable logistic regression models and ordinal risk index models, National Birth Defects Prevention Study, 1997–2007

Model	P-value from Hosmer-Lemeshow goodness of fit	Area under the curve (95% CI)
Multivariable models:		
Spina bifida	0.88	0.55 (0.50–0.60)
Anencephaly	0.91	0.59 (0.54–0.66)
Anencephaly + spina bifida	0.94	0.56 (0.53–0.60)
Ordinal risk index model:		
Spina bifida	0.71	0.55 (0.50–0.59)
Anencephaly	0.90	0.57 (0.52–0.63)
Anencephaly + spina bifida	0.72	0.56 (0.52–0.59)