Modeling behaviour of congenital anomalies diagnosed through prenatal USG based on risk factors

Prasun Das1* and Aiswaryya Deep Ghosh2
1Indian Statistical Institute, SQC & OR Division, 203, B T Road, Kolkata 108, India
2R G Kar Medical College and Hospital, Kolkata, India

Received 16 March 2009; accepted 19 October 2009

This study presents teratogenic factors and their deleterious effect on embryogenesis (congenital anomalies) in various age groups of pregnant mothers. A classification rule is developed based on significant risk factors in identifying possibility of future anomalies in mothers based on statistically significant risk factors. Developed system is suggested to using for early detection of cases, as a screening procedure and other preventive investigations as necessary.

Keywords: Classification rule, Congenital anomalies, Multivariate statistical analysis, Prenatal USG, Risk factors

Introduction

Human congenital anomalies had been reported1 in prehistoric cave paintings, sculptures and in writings, describing Talipes, Acondroplasia and conjoined twins etc. Most of the congenital anomalies are evident at birth while others may remain latent only to become clinically apparent at later phases of life2. Congenital anomalies are resultant of either genetic or environmental or, as in majority of cases, a combination of both, in their etiological origin3. Various factors causing congenital anomalies are: chromosomes, 6-7; mutant genes, 7-8; environmental factors, 7-10; multifactor Inheritance, 20-28; and unknown etiology, 50-60%. Special vulnerabilities arise owing to sensitivity of processes that occur only during early development4. Studies5 were carried out on understanding of teratogenic agents and their effects on developing organisms. A large number of studies had shown a positive association between advancing parental age and congenitally malformed babies6-9. Further, studies are available on the role of maternal obesity on congenital anomalies10-12. Diabetic women develop complications during pregnancy and postpartum when compared to women without diabetes2,18. Relation between maternal anaemia and birth weight has been reviewed extensively19,20. However, very little work is available on actual effect of maternal anaemia on congenital anomalies4. Studies15,16 had already shown ominous effects of maternal folic acid deficiencies in development of CNS (central nervous system) and orofacial region of embryo. Few studies were conducted on preconceptional use of folic acid, which proved that almost equal weightage should be put on this factor during pregnancy17,18. Women are placed in potentially compromising situations by exposure to drugs prior to diagnosis of pregnancy since at least half of all pregnancies are unplanned4. Most accepted classification of drugs according to their teratogenicity status is into categories A, B, C, D and X19. Loebstein20 observed that all decisions about drug use in pregnancy should be individualized. In heavy cigarette smoking, premature deliveries and congenital anomalies are considerably higher than in mothers who do not smoke. Types of anomalies, which are common in smoking mothers, are cardio-vascular and limb-defects including urinary tract anomalies. Resultant effect is fetal hypoxia and deficiency of important nutrients causing abnormal development2,21-22. Infants of mothers who consume large amount of alcohol during pregnancy may exhibit a pattern of physical anomalies referred to as fetal alcohol syndrome23,24.

This study presents effects of few selected environmental risk factors on developing embryo.

Materials and Methods

All expectant (pregnant) mothers of pregnancy between 18 to 20 weeks of gestational age, coming to
Fig. 1—Respiratory system defect: Congenital hyper-echoic lung

Fig. 2—Congenital limb defect
Fig. 3—CNS defect (encephalocele)

Fig. 4—CNS defect (sacro-coccygeal meningocele)
Fig. 5—Genito-urinary system defect (congenital hydronephrosis)

Fig. 6—Hugely distended bladder
radiology department for routine prenatal ultrasonography, were chosen as subjects for the study. All multiple pregnancies (twins, triplets etc.) are excluded from the present study. Mothers, with no fetal profile abnormality detected after ultrasonography, were chosen as ‘normal’ group whereas mothers having one or more abnormalities in fetus were chosen as ‘abnormal’ group. A total of 1513 pregnancies were collected during 1½ year. Out of these, 14 mothers were ultrasonographycally diagnosed to have congenitally ‘abnormal’ fetuses (congenitally malformed babies). Technologies of ultrasound (3.5 MHz and 5 MHz probe, like M-mode, colour doppler study etc.) were used as and when required for searching congenital anomalies. Information on 11 possible influential (causal) variables (risk factors), namely, Age of Mother (MA), Age of Father (PA), Body Mass Index (BMI), Maternal blood sugar fasting (BS-F), Maternal hemoglobin percentage (Hb%), Maternal folic acid intake (FA), Maternal drug intake status (DG), Maternal smoking (SM), Maternal alcohol consumption (AL), Consanguous marriage (marriage between blood related individuals, mostly cousins and first order relatives CM) and Order of birth (Number of previous children, living or dead, and abortions-OB) were gathered for each of 1513 subjects. All these risk factors are proved to have a higher potentiality of harming developing embryo maximally during its vulnerable period. Selection of certain risk factors was on the basis of overall low socio-economic status of the population under study. For statistical analysis, PA, BMI, BS-F, Hb% and OB were considered with their actual numerical values whereas for other risk factors, a befitting numerical representation was made as follows: FA (pre-conceptional, 2; in first trimester, 1; none, 0); DG [no drugs (except vitamins), 0; category A, 1; category B, 2; category C, 3; category D, 4; category X, 6 (on the basis of teratogenecity)]; SM [no smoking, 0; smoking, 1; heavy smoking (> 20 cigarettes/day), 3]; AL (none, 0; regularly, 1); and CM (no, 0; yes, 1). Ultrasonographic plates (Figs 1-6) of few selected cases of this study are presented. Statistical treatment of above multidimensional data was done using linear discriminant analysis.

**Results**

From the set of mothers having ‘normal’ fetuses, a representative sample of 100 mothers was drawn, by simple random sampling (SRS). Summary statistics (min, max, average) of explanatory variables for two groups were computed (Table 1). After initial scrutiny, presence of three risk factors (SM, AL and CM) was found extremely rare in database with no statistical significance. Thus their low level of presence in present population may be an extremely good prognostic indicator as far as pregnancy complicated by congenital anomalies is concerned. Consequently, for remaining variables (8), in order to detect whether there exists any significant difference in explanatory variables between ‘normal’ and ‘abnormal’ groups, data were analyzed using multivariate analysis of variance (MANOVA) after verifying underlying assumptions (normality and homoscedasticity). Three test statistics (Pillai, Hotelling and Wilk’s lambda tests) show that equivalent F-value is

### Table 1—Summary statistics

<table>
<thead>
<tr>
<th>Sl No.</th>
<th>Variable</th>
<th>Normal (n=100)</th>
<th>Abnormal (n=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>1</td>
<td>MA</td>
<td>19</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>PA</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>BMI</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>BS-F</td>
<td>80</td>
<td>112</td>
</tr>
<tr>
<td>5</td>
<td>Hb%</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>FA</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>DG</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>SM</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>AL</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>CM</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>OB</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

MA, Age of mother; PA, Age of father; BMI, Body Mass Index; BS-F, Maternal blood sugar fasting; Hb%, Maternal hemoglobin percentage; FA, Maternal folic acid intake; DG, Maternal drug intake status; SM, Maternal smoking; AL, Maternal alcohol consumption; CM, Consanguous marriage; OB, Order of birth
It was therefore concluded that explanatory variables of two groups were significantly different. After having found significant difference with respect to eight explanatory variables, nature of difference was explored using discriminant analysis in two stages. Classification functions of two groups after statistical calculation were obtained as

For Group “Normal”
\[ L_1 (\mathbf{x}) = -205.360 + 2.478 \times MA + 0.441 \times PA + 2.645 \times BMI + 1.808 \times BS-(F) + 11.053 \times Hb% - 5.316 \times FA - 7.997 \times DG - 7.222 \times OB \] …(1)

For Group “Abnormal”
\[ L_2 (\mathbf{x}) = -235.906 + 2.912 \times MA + 0.272 \times PA + 2.808 \times BMI + 2.169 \times BS-(F) + 8.176 \times Hb% - 2.925 \times FA - 7.909 \times DG - 7.434 \times OB \] …(2)

where, \( \mathbf{x} = [MA, PA, BMI, BS-(F), Hb\%, FA, DG, OB] \).

Thus, discriminant function between two groups was found as

\[ D_s (\mathbf{x}) = L_1 (\mathbf{x}) - L_2 (\mathbf{x}) = 30.546 - 0.434 \times MA + 0.169 \times PA - 0.163 \times BMI - 0.361 \times BS-(F) + 2.877 \times Hb\% - 8.241 \times FA - 0.088 \times DG + 0.212 \times OB \] …(3)

Canonical analysis indicated that discriminant function \( D_s (\mathbf{x}) \) is highly significant with a canonical correlation of \( R = 0.828 \) (Wilk’s lambda = 0.314448, Chi-square value = 124.9491 with 8 d.f. and \( p \)-value < 0.00001). Thus, about 68% of the between-group variability was accounted for by discriminant function. Standardized canonical discriminant function was computed as

\[ C_s (\mathbf{x}) = 0.308376 \times MA - 0.125833 \times PA + 0.089089 \times BMI + 0.544128 \times BS-(F) - 0.62617 \times Hb\% + 0.46386 \times FA + 0.005726 \times DG - 0.037486 \times OB \] …(4)

As indicated by absolute magnitude of standardized coefficients, Hb% had highest contribution in discriminating between two groups followed by BS-(F), FA and MA. Respective \( p \)-values showed highest significance level for Hb% \((p=0.00000001)\), BS-(F) \((p=0.000003)\), FA \((p=0.000122)\). MA, though having a higher \( p \)-value, had a very high tolerance level (tolerance = 0.835144).

In order to determine optimal classification rule, threshold value for discrimination between two groups, based on sample proportions of two groups, is estimated as -1.96432. Hence, optimal classification rule is as follows: Allocate a patient with his observation \( \mathbf{x} \) in “Normal” group, if \( D_s (\mathbf{x}) > -1.96432 \) and in “Abnormal” group, if \( D_s (\mathbf{x}) \leq -1.96432 \).

Using this optimal classification rule, sample observations were classified into two groups. Total cases classified correctly were found to be 97.4%. A total of three cases were misclassified; out of which, two cases...
[-55 (Mahalanobis distance (MD) from ‘normal’ group centroid = 30.20728 and from ‘abnormal’ group centroid = 17.83254; posterior probability = 0.014468 for ‘normal’ and 0.985532 for ‘abnormal’) and 98 (MD from ‘normal’ group centroid = 23.80386 and from ‘abnormal’ group centroid = 16.49744; posterior probabilities for ‘normal’ = 0.156158 and for ‘abnormal’ = 0.843842)] were misclassified as “abnormal” whereas case no. 102 (MD from ‘normal’ group centroid = 14.87131 and from ‘abnormal’ group centroid = 12.35130; posterior probabilities = 0.669541 for ‘normal’ and 0.330459 for ‘abnormal’) has been identified as “normal” patient. Canonical discriminant scores of 114 cases were obtained by evaluating unstandardized canonical discriminant function at the values of explanatory variables for each case. All scores are positive for “abnormal” case with (min, max, average) = (1.80948, 6.26318, 3.89216) where minimum occurs at observation no. 102, which is a case of misclassification (Fig. 7), because abnormality is not at all severe. Corresponding Mahalanobis distance (Fig. 8) and posterior probability are also very much insignificant to belong in ‘abnormal’ group. However, for “normal” cases, small scores with both positive and negative signs are observed except the cases 55 and 98, where values are high compared to other cases in “normal” group.

Discussion

Present study is based on prenatal diagnosis of malformed babies in Indian perspective. Firstly, in India all women coming to hospitals may not end up delivering there. Some such babies may be born dead or spontaneously aborted and their malformed nature might not be documented or even identified in peripheral centers and in-house deliveries. Thus there is a high probability of missing a lot of actual cases. Hence, present approach ensures a higher probability of detecting malformed cases. Overall sensitivity to detect all fetal abnormalities, during 18 to 20 weeks of gestational age, during routine screening for normal fetal parameters, were found to be more than 70%. However, when such procedure is done by expert radiologists with a prime motive to detect fetal abnormalities, as done in the present study, sensitivity of detection is more27.

Discriminant analysis showed three cases of misclassification [case numbers 55, 98 (normal) and 102 (abnormal)]. Case number 55 showed making it closer to abnormal group. Most of the major risk factors (MA, 33; PA, 38; BMI, 28; BS-F, 112; Hb%, 9;OB, 2) show extreme value and rest are within normal range. So, this is a very rare possibility to be ‘normal’. But case number 98 is a true misclassification as almost all risk factors have a normal value. Moreover, it is placed in abnormal group only on the basis of posterior probability, which has perhaps a higher significance than Mahalanobis distance in prediction accuracy. Such misclassifications are explained among the cases which are represented by sample observations only from ‘normal’ group of mothers. Case number 102 had a relatively minor abnormality of limb defects. Moreover,
Mahalanobis scores of this case of normal (14) and abnormal (12) are very close, showing a borderline situation and hence discrimination rule is quite justified.

Advancing maternal age was found to be a very important risk factor (Wilks’ lambda 0.317865, \( p \)-level 0.287944) amongst the first four statistically important risk factors, even with relatively high \( p \)-value. Proportion of ‘normal’ mothers is observed lower than that of mothers with abnormal babies over age distribution, except at a higher age. Average age of motherhood is going up with rapid growth of urbanization, which leads to a much high probability of having deleterious effect on pregnancy especially in presence of other risk factors. However, from present study, no statistically significant correlation could be obtained regarding types of anomalies in various age groups diagnosed of mothers. Advancing paternal age had also been marked as a statistically significant risk factor (Wilks’ lambda 0.315070, \( p \)-level 0.649522). Like maternal age, paternal age also follows a similar pattern showing older age distributions in case of fathers of abnormal babies. Though, present study is unable in specifying exact defects caused by advancing paternal age but it does definitely show that increased paternal age has a positive association with congenital anomalies. Maternal BMI showed a moderately high association with congenital anomalies (Wilks’ lambda 0.315953, \( p \)-level 0.479945). Like maternal age, paternal age also follows a similar pattern showing older age distributions in case of fathers of abnormal babies. Though, present study is unable in specifying exact defects caused by advancing paternal age but it does definitely show that increased paternal age has a positive association with congenital anomalies. Maternal BMI showed a moderately high association with congenital anomalies (Wilks’ lambda 0.315953, \( p \)-level 0.479945). Maternal obesity is a growing problem even in oriental population. Present work is comparable to most of earlier studies, showing that congenital anomalies are positively associated with increased BMI. But unlike in most of these works, in which BMI beyond 24.9 is taken to be overweight and beyond 29.9 as obesity, present work deals with BMI as a continuous variable. Similar results had been obtained with blood sugar (fasting) level, which also shows a very strong positive correlation (Wilks’ lambda 0.386438, \( p \)-level 0.000003). Blood sugar fasting was chosen over postprandial sugar as it gives a more specific measure of gestational glycaemic status as individual dietary factors are minimized. Moreover, a presumptive diagnosis of overt diabetes is done on the basis of fasting blood sugar level (>105 mg/dl). In present study, this explanatory variable turned out to be very indicative and hence was a part of classification rule. A very high association of anaemia has been observed with congenital anomalies (Wilks’ lambda 0.423902, \( p \)-level 0.000001). Results of present study differ from most of the studies worldwide, since anaemia is not as prevalent in west as in India. Maternal folic acid intake status turned out to be a vitally important risk factor (Wilks lambda 0.362160, \( p \)-level 0.000122). Both preconception and post conceptional folic acid intake status was taken with equal weightage for relative importance of fetal development. This finding is strongly comparable with other important studies worldwide. This study shows a very low discrimination power of drugs (Wilks lambda 0.314455, \( p \)-level 0.963729), signifying that there is a low prevalence of teratogenic drug use in the present population. This might be due to low access to drugs and also tolerance of minor illness without using drugs in Indian population than in the West. Drug abuse is also lower. However, more detrimental drugs of category C and category D are usually used on severely ill, in-house mothers, who are not properly represented in the present study. Parity is an important risk factor as congenital anomalies are somewhat related to ageing of ovum and decrease in maternal nutritional reserves due to multiple births.

Conclusions

Out of 11 risk factors playing their combined role towards formation of congenital anomalies, four factors (maternal age, blood sugar level, folic acid and maternal haemoglobin percentage) turned out to be extremely important in Indian perspective. These factors are at par with similar studies worldwide barring haemoglobin percentage. Construction of unique classification rule in this study, deduced as a function of significant risk factors, could predict future chances of a normal or an abnormal baby by only studying risk factors of the mother. Such type of early prediction gives an excellent opportunity not only for early diagnosis but also for early management of cases. It might also be used as a screening procedure to detect probable abnormal cases. This process has thus ensured maximum number of prenatal case identification in Indian perspective, where birth record system is underdeveloped to identify congenital anomalies.

References

1. Sadler T W, Langman’s Medical Embryology, 10th edn, 84 (Lippincott Williams & Wilkins, Philadelphia) 2006, 67-121.


