



Detection of Risk Factor Interactions in Cerebral Palsy Using Statistical Methods

Enayatollah Bakhshi,¹ Sepideh Zare Delavar,¹ Farin Soleimani,² and Akbar Biglarian^{1,*}

¹Department of Biostatistics, Social Determinants of Health Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

²Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

*Corresponding author: Akbar Biglarian, Department of Biostatistics, Social Determinants of Health Research Center, University of Social Welfare and Rehabilitation Sciences, Koodakyar Ave, Evin, Tehran, Iran. Tel: +98-2122180146, E-mail: abiglarian@gmail.com

Received 2017 July 01; Revised 2018 February 13; Accepted 2018 February 23.

Abstract

Background: The interaction of the risk factors is an important focus in medical study. The etiology of CP is multi-factorial and usually arises from a series of causal pathways.

Objectives: The aim of this study was to compare four statistical methods to detect the interactions of risk factors of cerebral palsy (CP).

Methods: In this cross-sectional study, 225 children aged 1 - 6 years were studied during 2008 and 2009 in Tehran, Iran. Penalized logistic regression (PLR), multifactor dimensionality reduction (MDR), logic regression (LR) and classification regression (CR) methods were used to detect interactions.

Results: There was an interaction between the consanguinity factor and gender, asphyxia, preterm delivery, vaginal delivery, previous pregnancies, history of pregnancy, history of illness, sepsis, and small for gestational age (SGA). The methods used in this study were detect two-way (6 terms), three-way (11 terms) and four-way (2 terms) interaction terms. Of these interactions, asphyxia × consanguinity interaction was detected by MDR, LR, and CR methods.

Conclusions: Our findings showed that the consanguinity factor had an interaction with most of the risk factors and especially consanguinity × asphyxia term was detected in all methods. So this interaction term is very important for considering by clinicians.

Keywords: Penalized Logistic Regression, Multifactor Dimensionality Reduction, Logic Regression, Classification Regression, Cerebral Palsy

1. Background

Cerebral palsy (CP) consists of a heterogeneous group of non-progressive clinical syndromes that is caused by abnormal development of the brain and is the most common cause of physical disability in children (1-4). Despite medical and technological progresses in neonatal and prenatal intensive care in the past two decades, movement, and/or cognitive and/or behavioral disorders are still a dilemma in children (5). In general population, the prevalence of CP was reported 1.5 to 2.5 per 1000 live births (6-8). The etiology of CP is multi-factorial (9) and various, and usually arises from a series of causal pathways (7, 10).

Some risk factors for cerebral palsy were identified by investigators such as: low birth weight, premature birth, multiple births, infections during pregnancy, medical conditions of the mother, birth complications, brain infections, injury, asphyxia, preterm delivery, and multiple fetuses (7, 10-12). It is mentioned that having a risk factor does

not mean that an infant will have CP (3, 4), but the presence of some risk factors will increase the chance of it (9, 10), i.e., more often multiple risk factors lead to CP (10).

In the most of studies, the risk factors of CP were recognized and reported but the interaction terms of those risk factors were not recognized. The interaction terms effects are of consequence for doctors and therapists and so identification of these terms is of great importance.

Regression analysis provides a way to identify the interaction terms between the risk factors. In the case of a binary response variable, Logistic regression can be used to explain the relationship between the risk factors and response variable (13). But when the independent variables are too large, logistic regression is not efficient to identify interactions. In such situations, penalized logistic regression (PLR), multifactor dimensionality reduction (MDR), Logic regression (LR) approach and classification regression (CR) methods can be used (13-17). This study, the first known study in Iran, adds to the existing information and

highlights interaction between main factors. The main purpose of this study is to detect interaction terms effects for risk factors of CP children by PLR, MDR, LFS, and CR methods.

2. Methods

2.1. Data Source

This was a cross-sectional study in which the data of 225 children aged 1 - 6 years was used. This study was approved by university of social welfare and rehabilitation sciences' ethics committee (code IR.USWR.REC.1394.190). All data was gathered over 12 months, March 2008 to February 2009, in Tehran. 112 subjects with documented CP entered the study. These children were referred from health care centers, located in the eastern and northern parts of the city of Tehran, to Asma comprehensive rehabilitation center as a referral center for the city of Tehran. In addition 113 subjects, who were apparently healthy, i.e. without CP, were randomly selected. "These children had attended only for well-being check-ups (as the control group), and were examined in the same health-care centers by the research team" (5).

2.2. Variables

These related factors of perinatal and neonatal, as the risk factors of CP in children, were collected and used: gender, breech birth, neonatal respiratory disorders, jaundice, sepsis, history of illness, parental consanguinity, premature rupture of membrane (PROM), history of infertility, multiple birth in a pregnancy, previous pregnancies, vaginal bleeding, asphyxia, SGA, preterm delivery (gestational age 37 weeks or less), risky delivery (the first delivery or more than 4 deliveries), history of abortion, vaginal delivery or cesarean delivery, maternal age at delivery (risks: < 16 or > 40 years).

2.3. Statistical Methods

PLR, MDR, LR and CR methods were used to detect interactions. Data analysis was carried out with the R3.2.2 software.

2.3.1. Penalized Logistic Regression

PLR can be considered as generalized regression model which finds interactions with the large number of variable. In this case, the logistic regression criterion is used by combining with a penalization of the L2-norm of the coefficients. "The quadratic penalty makes it possible to code each level of a factor by a dummy variable, yielding coefficients with direct interpretations. This coding method cannot be applied to regular logistic regression because

the dummy variables representing a factor are perfectly collinear" (13). In this method, because of avoiding overfitting, a penalty on large fluctuations is imposed in the process of parameters estimation:

$$Q = \frac{1}{n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^m x_{ij} \beta_j \right)^2 + \lambda J(\beta) \quad (1)$$

where y is a vector of dependent variables; x is a matrix of independent variables; β is a $m \times 1$ matrix of parameters; λ is tuning parameter; and $J(\beta)$ is penalty term (13, 17).

2.3.2. Multifactor Dimensionality Reduction

The MDR is a nonparametric method, as an alternative to logistic regression, to detect the interactions of main effects even if the main effects are statistically insignificant. So, this is a weakness of this approach. However, this method considers all combinations of potentially interaction terms (16) and will be able to detect high-order interaction terms (such as two-way and/or three-way interactions, etc.) (17, 18). The MDR method creates a measure of accuracy and then the best model is selected based on the highest accuracy. It is expected that this model perform well in terms of prediction and assessment of internal validation measures (19, 20). For comparing different interactions terms (combination of variables), balanced accuracy (BA) is used as

$$BA = \frac{1}{2} \left(\frac{TP}{FN + TP} + \frac{TN}{FP + TN} \right) \quad (2)$$

where FP, FN, TP, TN represent the number of false positives, false negatives, true positives, and true negatives, respectively. Indeed BA is the arithmetic mean of sensitivity and specificity (21).

2.3.3. Logic Regression

LR is a (generalized) regression methodology that is applied to detect interaction terms between covariates, especially, when most/all of the covariates are binary. In the LR, we are interested in finding the binary variables which resulting Boolean combinations of initial binary covariates. These new covariates, new combinations, will be entering in the regression model to get the best fit. The LR model is defined as

$$g(E(Y)) = \beta_0 + \sum_{j=1}^t \beta_j L_j \quad (3)$$

where Y is response variable, β_0 and β_j are parameters, and L_j is Boolean combination (logic term) of the covariates. In this model, logic terms must be determined and parameters must be estimated simultaneously (14). In this case, a single-tree approach of logic regression was used to detect interaction terms. A large value of variable importance measure (VIM) means a high importance of a particular interaction, whereas a value of about zero means no importance for it (22).

2.3.4. Classification Regression Methods

CR method is one of the most powerful, non-parametric, and yet simple powerful statistical method to determine the most important covariates, and to visualize the important associations and also to find the accurate prediction of outcome (15). Methods which will be used in this sense are: classification and regression tree (CART), Adaboost, Bagging, and C 4.5 algorithm. These methods, also, will detect the existing interaction between covariates.

3. Results

The average weight of the children at birth was 2491.94 ± 887.13 and 3101.77 ± 542.83 with medians of 2500 and 3150 grams for CP and non CP group respectively. The average gestational age was 35.66 ± 3.78 and 39.81 ± 1.32 with medians of 38 and 40 weeks for CP and non CP group respectively. 47.1% of the studied children were male. In total, 14.7% of children presented with respiratory disorders, 24.4% with jaundice, 14.7% with sepsis, 17.8% with history of illness, 69.8% with consanguinity, 9.3% with PROM, 4.9% with adjuvant treatment of infertility, 4.9% with multiple birth in a pregnancy, 7.1% vaginal bleeding, 50.2% with asphyxia, 20.9% with SGA, 19.6% with preterm delivery, 1.3% with high risk delivery, 17.3% with a history of abortion, 19.6% with previous pregnancy, 70.7% with vaginal delivery and 7.6% of maternal risky age (Table 1). The relation between these risk factors and response variable was examined by chi-square and Fisher exact tests. The results of the chi-square test showed that neonatal respiratory disorders ($P < 0.001$), jaundice ($P = 0.040$), neonatal infant ($P < 0.001$), history of illness ($P = 0.001$), history of consanguinity ($P < 0.001$), PROM ($P = 0.003$), adjuvant treatment of infertility ($P = 0.028$), multiple birth in a pregnancy ($P = 0.001$), asphyxia ($P < 0.001$), preterm delivery ($P < 0.001$), have a significant relationship with CP whereas with ordinary logistic regression analysis, only history of illness ($P = 0.002$ and $OR = 0.03$) and asphyxia ($P < 0.001$ and $OR = 127.579$) variables were significant.

For identifying the interactions by PLR, first, the optimal model based on the BIC criterion and with different values of (tuning parameter) was evaluated. Then, with smallest deviation criterion was considered as a tuning parameter. The value was equal to 0.005. The identified main and interaction effects, from PLR model with forward stepwise and forward stage-wise procedure, are reported in Table 2. The main effects of asphyxia ($P < 0.001$), consanguinity ($P = 0.003$), preterm delivery ($P = 0.002$), history of illness ($P = 0.001$) were significant effects on CP. In the forward stage-wise procedure, the main effects of asphyxia ($P < 0.001$), consanguinity ($P < 0.001$) and also interaction

Table 1. Characteristics of Children with and Without CP, Based on Their Risk Factors^a

Variables	Control, N = 113	Case, N = 112	P Value ^b
Male gender	47 (20.9)	59 (26.2)	0.125
Breech birth	4 (1.8)	4 (1.8)	1.000
Respiratory disorders	2 (0.9)	31 (13.8)	< 0.001
Jaundice	21 (9.3)	34 (15.1)	0.058
Sepsis	5 (2.2)	28 (12.4)	< 0.001
History of illness	30 (13.3)	10 (4.4)	0.001
Consanguinity	50 (22.2)	107 (47.6)	< 0.001
PROM	4 (1.8)	17 (7.6)	0.006
Adjuvant treatment of infertility	2 (0.9)	9 (4.0)	0.061
Multiple birth	0 (0.0)	11 (4.9)	0.002
Vaginal bleeding	6 (2.7)	10 (4.4)	0.426
Asphyxia	23 (10.2)	90 (40.0)	< 0.001
SGA	19 (8.4)	28 (12.4)	0.178
Preterm delivery	3 (1.3)	41 (18.2)	< 0.001
High risk delivery	0 (0.0)	3 (1.3)	0.242
History of abortion	17 (7.6)	22 (9.8)	0.462
Previous pregnancy	17 (7.6)	27 (12.0)	0.122
Vaginal delivery	64 (28.4)	95 (42.2)	< 0.001
Maternal age (< 16 or > 40 y)	6 (2.7)	11 (4.9)	0.304

^aValues are expressed as No. (%).

^b2-sample test for equality of proportions with continuity correction.

terms of preterm delivery \times consanguinity ($P = 0.040$), history of illness \times consanguinity ($P < 0.002$) had significant relationship with CP.

For analyzing the data with MDR, cross validation and 3-way split methods were used to identify the more important compounds.

The results are summarized in Table 3. Both methods showed that the interaction of asphyxia \times consanguinity \times preterm delivery had the most accuracy, and then the MDR procedure with three-way interaction was fitted to the data. Also, Table 3 shows the identified interaction terms based on balance accuracy. The first compound, i.e. asphyxia \times consanguinity \times preterm delivery, has higher balance accuracy and hence was detected as an important interaction term. This interaction means that the effect of consanguinity on CP is not equal in the different levels of asphyxia \times preterm delivery.

In LR approach, the most logical important compounds on CP based on VIM_{single} index were reported. In this case, Asphyxia \times consanguinity \times history of illness compound ($VIM = 8.00$) was most important logi-

Table 2. Estimated Penalized Logistic Regression Coefficients with Forward Step-Wise and Stage-Wise Procedure

Variables	Coefficient	SE	P Value
Intercept ^a	9.58	2.61	< 0.001
Asphyxia	4.28	0.71	< 0.001
Consanguinity	7.36	2.48	0.030
Preterm delivery	7.59	2.46	0.020
History of illness	-2.95	0.90	0.010
High risk delivery	5.14	3.78	0.174
Intercept ^b	6.21	1.19	< 0.001
Asphyxia	3.50	0.59	< 0.001
Consanguinity	4.57	1.12	< 0.001
Preterm delivery × consanguinity	6.52	3.17	0.040
History of illness × consanguinity	-2.75	0.88	0.002
Gender × preterm delivery × consanguinity	1.65	3.74	0.660

^a Forward step-wise procedure.^b Forward stage-wise procedure.

cal compound. Other identified logical compounds were consanguinity × preterm delivery (VIM = 1.70), history of illness × consanguinity × PROM × asphyxia (VIM = 1.60), consanguinity × asphyxia (VIM = 1.55), consanguinity × PROM × asphyxia (VIM = 0.55), PROM × SGA (VIM = 0.55).

In addition, the result of stepwise logistic regression analysis based on the risk factors and detected logical compounds showed that two logical compounds (history of illness × consanguinity × PROM × asphyxia, and consanguinity × PROM × asphyxia) were significant. For completing the analysis, the significant variables in the ordinary logistic regression without any interactions and these two logical compounds, were modeled and asphyxia variable with two mentioned compounds ($P < 0.001$) remained in the model.

Finally, the consequences of the AdaBoost and Bagging methods, C4.5 algorithm, and CART are reported in Table 4. Then the performance of these methods with four criteria (sensitivity, specificity, and error as well as Matthew's correlation coefficient (MCC)), were compared with the 10 stages cross-validation method. The AdaBoost method with lower error (0.06) and higher specificity and sensitivity (0.94 and 0.95, respectively) had a better performance than other methods (Table 5).

4. Discussion

In many epidemiological studies to investigate the relation of risk factors with a special case like presence or absence of a disease, as a dependent variable, logistic regres-

sion model was used. If the numbers of dependent variables were large, determining and identifying the interaction becomes difficult or even impossible. In this study, PLR, MDR, LR, and CR models were used to detect interaction terms.

In several studies, improvement of these models has been reported. Park and Hastie have used the PLR and MDR models to determine interaction, gene-gene and gene-environment, terms in blood pressure and Ladder cancer data (13). Sun and Wang have proposed the PLR model for genetic data and showed that this model can be used when data is correlated within a group (18). Stoknes et al. in a study on the Norwegian children with CP disease have identified several interaction terms (preterm birth × induction, maternal disease × preterm birth, maternal disease × premature birth, maternal disease × induction, maternal disease × low 5-minutes Apgar score) between related risk factors. The only significant interaction was reported as maternal disease × preterm birth. In addition, they reported that when the number of risk factors increases, the risk for CP will increase (23). O'Callaghan et al. have not found interaction between SNPs and epidemiologic risk factors for CP outcome. Finally, they have suggested more studies to assess interaction between them (24). Gao et al. in a meta-analysis study showed that maternal age (≥ 35 years), multiple pregnancy, and medicine use in early pregnancy, harmful environment, recurrent vaginal bleeding during pregnancy and pregnancy-induced hypertension were the risk factors for CP (25). Consanguinity was reported as major risk factor in several studies in southwest Asia (26-29). Studies showed that the consanguinity was associated with complex disorders (27) and increases susceptibility to multifactorial diseases (28). In addition, it is considered as a factor of higher reading disabilities (29, 30) and the cause of developmental problems (31). They have not reported any interaction terms between risk factors.

In most studies of CP no interaction between the risk factors was reported. Models without any interactions occur, when the number of risk factors is large and so determining the interactions will be difficult or even impossible. However, based on used method in this study, consanguinity, preterm birth, and asphyxia have had the most interaction with other risk factors. In this scene, consanguinity is a cultural factor and also important to investigate. One of the limitations of this study was the sample size in CP group. The cultural, local and available health system, e.g. NICU can probably affect the proportion of consanguinity and asphyxia and also homogeneity of other risk factors in this sample. So the researchers should pay attention to this issue, conduct their research using larger sample size in other CP population studies.

Table 3. The MDR Fit Using Three-Way Split and Cross Validation Methods

Number of Interactions	Final Model with 3-Way Split	Training Accuracy	Testing Accuracy	Validation Accuracy
k = 4	Asphyxia × consanguinity × preterm delivery	97.83	91.07	82.45
k = 3	Asphyxia × consanguinity	87.82	91.96	85.83
k = 2	Asphyxia × consanguinity	86.43	92.31	90.38
Number of interactions	Final model with cross validation	Classification accuracy	Prediction accuracy	
k = 4	Asphyxia × consanguinity × preterm delivery	92.13	91.92	
k = 3	Asphyxia × consanguinity × preterm delivery	92.15	92.28	
k = 2	Asphyxia × consanguinity	89.32	89.03	
Top 5 models identified by MDR method with the best testing accuracy				Balance accuracy
Asphyxia × consanguinity × preterm delivery				92.145
Asphyxia × consanguinity × respiratory disorders				90.459
Sepsis × asphyxia × consanguinity				90.001
Consanguinity × PROM × asphyxia				90.001
Consanguinity × multiple birth × asphyxia				89.897

Table 4. Detected Interactions by AdaBoost, Bagging, C4.5 Algorithm, CART Methods^a

Method	Interactions
Adaboost	
2 way	Consanguinity × gender Previous pregnancy × vaginal delivery
3 way	Consanguinity × preterm delivery × gender Preterm delivery × history of illness × asphyxia Consanguinity × asphyxia × gender History of illness × gender × SGA Sepsis × asphyxia × SGA History of illness × asphyxia × gender Preterm delivery × vaginal delivery × asphyxia
Bagging	
2 way	Consanguinity × asphyxia
3 way	Consanguinity × preterm delivery × asphyxia
C4.5 algorithm	
2 way	Asphyxia × preterm delivery
4 way	Asphyxia × consanguinity × history of illness × preterm delivery
CART	
2 way	Asphyxia × consanguinity

Source: Based on data in Zare Delavar S, IJE 2014; 10(2): 33-9.

4.1. Conclusions

Detection and identification of some interaction terms between risk factors of CP using PLR, MDR, LR and CR models were most important results of this study and specially, there was an interaction between the consanguinity factor and other factors. These models have played an important role in the identification of interaction terms in real data (16-18, 22, 32) and simulation data set (14, 17, 22, 33). Such statistical methods have the ability of detecting risk factors interactions of diseases, such as CP. It is mentioned that this study is the beginning for a comprehensive discussion on interaction terms between risk factors in CP patients. Our results can be used by therapists and clinicians to design a preventive strategy for these patients.

Acknowledgments

The authors are thankful to all health care centers, affiliated to Shahid Beheshti University of Medical Sciences, managers and personnel and also all families who cooperated in performing this study. In addition, the authors wish to express their gratitude to the deputy of research and technology of University of social welfare and rehabilitation Sciences for financial support.

Table 5. Sensitivity, Specificity, and Error in AdaBoost, Bagging, C4.5 Algorithm, and CART Methods^a

Method	Sensitivity	Specificity	Error	MCC
AdaBoost	0.941 ± 0.029	0.951 ± 0.030	0.06 ± 0.006	0.887 ± 0.028
Bagging	0.832 ± 0.027	0.909 ± 0.029	0.135 ± 0.007	0.737 ± 0.025
C4.5 algorithm	0.892 ± 0.030	0.927 ± 0.029	0.095 ± 0.009	0.813 ± 0.029
CART	0.856 ± 0.030	0.894 ± 0.028	0.115 ± 0.010	0.769 ± 0.030

Source: Based on data in Zare Delavar S, IJE 2014;10(2): 33-9.

Footnote

Funding/Support: This study was supported by the deputy of research and technology of University of social welfare and rehabilitation Sciences.

References

- Tollanes MC, Wilcox AJ, Lie RT, Moster D. Familial risk of cerebral palsy: population based cohort study. *BMJ*. 2014;**349**:g4294. doi: [10.1136/bmj.g4294](https://doi.org/10.1136/bmj.g4294). [PubMed: [25028249](https://pubmed.ncbi.nlm.nih.gov/25028249/)]. [PubMed Central: [PMC4099475](https://pubmed.ncbi.nlm.nih.gov/PMC4099475/)].
- Rackauskaite G, Uldall PW, Bech BH, Ostergaard JR. Impact of child and family characteristics on cerebral palsy treatment. *Dev Med Child Neurol*. 2015;**57**(10):948-54. doi: [10.1111/dmcn.12791](https://doi.org/10.1111/dmcn.12791). [PubMed: [25952577](https://pubmed.ncbi.nlm.nih.gov/25952577/)].
- Doyle LW, Crowther CA, Middleton P, Marret S. Antenatal magnesium sulfate and neurologic outcome in preterm infants: a systematic review. *Obstet Gynecol*. 2009;**113**(6):1327-33. doi: [10.1097/AOG.0b013e3181a60495](https://doi.org/10.1097/AOG.0b013e3181a60495). [PubMed: [19461430](https://pubmed.ncbi.nlm.nih.gov/19461430/)].
- Soleimani F, Vameghi R, Biglarian A. Antenatal and intrapartum risk factors for cerebral palsy in term and near-term newborns. *Arch Iran Med*. 2013;**16**(4):213-6. [PubMed: [23496363](https://pubmed.ncbi.nlm.nih.gov/23496363/)].
- Soleimani F, Vameghi R, Biglarian A, Daneshmandan N. Risk factors associated with cerebral palsy in children born in Eastern and Northern districts of Tehran. *Iran Red Crescent Med J*. 2010;**12**(4):428-33.
- Paneth N, Hong T, Korzeniewski S. The descriptive epidemiology of cerebral palsy. *Clin Perinatol*. 2006;**33**(2):251-67. doi: [10.1016/j.clp.2006.03.011](https://doi.org/10.1016/j.clp.2006.03.011). [PubMed: [16765723](https://pubmed.ncbi.nlm.nih.gov/16765723/)].
- Blair E, Watson L. Epidemiology of cerebral palsy. *Semin Fetal Neonatal Med*. 2006;**11**(2):117-25. doi: [10.1016/j.siny.2005.10.010](https://doi.org/10.1016/j.siny.2005.10.010). [PubMed: [16338186](https://pubmed.ncbi.nlm.nih.gov/16338186/)].
- Vincer MJ, Allen AC, Allen VM, Baskett TF, O'Connell CM. Trends in the prevalence of cerebral palsy among very preterm infants (<31 weeks' gestational age). *Paediatr Child Health*. 2014;**19**(4):185-9. doi: [10.1093/pch/19.4.185](https://doi.org/10.1093/pch/19.4.185). [PubMed: [24855414](https://pubmed.ncbi.nlm.nih.gov/24855414/)]. [PubMed Central: [PMC4028643](https://pubmed.ncbi.nlm.nih.gov/PMC4028643/)].
- Iida K, Takashima S, Takeuchi Y, Ohno T, Ueda K. Neuropathologic study of newborns with prenatal-onset leukomalacia. *Pediatr Neurol*. 1993;**9**(1):45-8. doi: [10.1016/0887-8994\(93\)90009-2](https://doi.org/10.1016/0887-8994(93)90009-2). [PubMed: [8452600](https://pubmed.ncbi.nlm.nih.gov/8452600/)].
- Nelson KB. Causative factors in cerebral palsy. *Clin Obstet Gynecol*. 2008;**51**(4):749-62. doi: [10.1097/GRF.0b013e318187087c](https://doi.org/10.1097/GRF.0b013e318187087c). [PubMed: [18981800](https://pubmed.ncbi.nlm.nih.gov/18981800/)].
- Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *N Engl J Med*. 2008;**359**(3):262-73. doi: [10.1056/NEJMoa0706475](https://doi.org/10.1056/NEJMoa0706475). [PubMed: [18635431](https://pubmed.ncbi.nlm.nih.gov/18635431/)].
- McIntyre S, Taitz D, Keogh J, Goldsmith S, Badawi N, Blair E. A systematic review of risk factors for cerebral palsy in children born at term in developed countries. *Dev Med Child Neurol*. 2013;**55**(6):499-508. doi: [10.1111/dmcn.12017](https://doi.org/10.1111/dmcn.12017). [PubMed: [23181910](https://pubmed.ncbi.nlm.nih.gov/23181910/)].
- Park MY, Hastie T. Penalized logistic regression for detecting gene interactions. *Biostatistics*. 2008;**9**(1):30-50. doi: [10.1093/biostatistics/kxm010](https://doi.org/10.1093/biostatistics/kxm010). [PubMed: [17429103](https://pubmed.ncbi.nlm.nih.gov/17429103/)].
- Ruczinski I, Kooperberg C, LeBlanc M. Logic regression. *J Comput Graph Stat*. 2003;**12**(3):475-511. doi: [10.1198/1061860032238](https://doi.org/10.1198/1061860032238).
- Breiman L, Friedman JH, Olshen RA, Stone CI. *Classification and regression trees*. Belmont, Calif: Wadsworth; 1984.
- Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet*. 2001;**69**(1):138-47. doi: [10.1086/321276](https://doi.org/10.1086/321276). [PubMed: [11404819](https://pubmed.ncbi.nlm.nih.gov/11404819/)]. [PubMed Central: [PMC1226028](https://pubmed.ncbi.nlm.nih.gov/PMC1226028/)].
- He H, Oetting WS, Brott MJ, Basu S. Power of multifactor dimensionality reduction and penalized logistic regression for detecting gene-gene interaction in a case-control study. *BMC Med Genet*. 2009;**10**:127. doi: [10.1186/1471-2350-10-127](https://doi.org/10.1186/1471-2350-10-127). [PubMed: [19961594](https://pubmed.ncbi.nlm.nih.gov/19961594/)]. [PubMed Central: [PMC2800840](https://pubmed.ncbi.nlm.nih.gov/PMC2800840/)].
- Sun H, Wang S. Penalized logistic regression for high-dimensional DNA methylation data with case-control studies. *Bioinformatics*. 2012;**28**(10):1368-75. doi: [10.1093/bioinformatics/bts145](https://doi.org/10.1093/bioinformatics/bts145). [PubMed: [22467913](https://pubmed.ncbi.nlm.nih.gov/22467913/)]. [PubMed Central: [PMC3348559](https://pubmed.ncbi.nlm.nih.gov/PMC3348559/)].
- Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, et al. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J Theor Biol*. 2006;**241**(2):252-61. doi: [10.1016/j.jtbi.2005.11.036](https://doi.org/10.1016/j.jtbi.2005.11.036). [PubMed: [16457852](https://pubmed.ncbi.nlm.nih.gov/16457852/)].
- Motsinger AA, Ritchie MD. The effect of reduction in cross-validation intervals on the performance of multifactor dimensionality reduction. *Genet Epidemiol*. 2006;**30**(6):546-55. doi: [10.1002/gepi.20166](https://doi.org/10.1002/gepi.20166). [PubMed: [16800004](https://pubmed.ncbi.nlm.nih.gov/16800004/)].
- Velez DR, White BC, Motsinger AA, Bush WS, Ritchie MD, Williams SM, et al. A balanced accuracy function for epistasis modeling in imbalanced datasets using multifactor dimensionality reduction. *Genet Epidemiol*. 2007;**31**(4):306-15. doi: [10.1002/gepi.20211](https://doi.org/10.1002/gepi.20211). [PubMed: [17323372](https://pubmed.ncbi.nlm.nih.gov/17323372/)].
- Schwender H, Ickstadt K. Identification of SNP interactions using logic regression. *Biostatistics*. 2008;**9**(1):187-98. doi: [10.1093/biostatistics/kxm024](https://doi.org/10.1093/biostatistics/kxm024). [PubMed: [17578898](https://pubmed.ncbi.nlm.nih.gov/17578898/)].
- Stoknes M, Andersen GL, Elkamil AI, Irgens LM, Skranes J, Salvesen KA, et al. The effects of multiple pre- and perinatal risk factors on the occurrence of cerebral palsy. A Norwegian register based study. *Eur J Paediatr Neurol*. 2012;**16**(1):56-63. doi: [10.1016/j.ejpn.2011.10.004](https://doi.org/10.1016/j.ejpn.2011.10.004). [PubMed: [22104566](https://pubmed.ncbi.nlm.nih.gov/22104566/)].
- O'Callaghan ME, Maclennan AH, Gibson CS, McMichael GL, Haan EA, Broadbent JL, et al. Fetal and maternal candidate single nucleotide polymorphism associations with cerebral palsy: a case-control study. *Pediatrics*. 2012;**129**(2):e414-23. doi: [10.1542/peds.2011-0739](https://doi.org/10.1542/peds.2011-0739). [PubMed: [22291124](https://pubmed.ncbi.nlm.nih.gov/22291124/)].
- Gao J, Zhao B, He L, Sun M, Yu X, Wang L. Risk of cerebral palsy in Chinese children: A N:M matched case control study. *J Paediatr Child Health*. 2017;**53**(5):464-9. doi: [10.1111/jpc.13479](https://doi.org/10.1111/jpc.13479). [PubMed: [28134474](https://pubmed.ncbi.nlm.nih.gov/28134474/)].

26. Erkin G, Delialioglu SU, Ozel S, Culha C, Sirzai H. Risk factors and clinical profiles in Turkish children with cerebral palsy: analysis of 625 cases. *Int J Rehabil Res*. 2008;**31**(1):89-91. doi: [10.1097/MRR.0b013e3282f45225](https://doi.org/10.1097/MRR.0b013e3282f45225). [PubMed: [18277211](https://pubmed.ncbi.nlm.nih.gov/18277211/)].
27. Masri A, Hamamy H, Khreisat A. Profile of developmental delay in children under five years of age in a highly consanguineous community: a hospital-based study-Jordan. *Brain Dev*. 2011;**33**(10):810-5. doi: [10.1016/j.braindev.2010.12.002](https://doi.org/10.1016/j.braindev.2010.12.002). [PubMed: [21194858](https://pubmed.ncbi.nlm.nih.gov/21194858/)].
28. Hamamy H, Antonarakis SE, Cavalli-Sforza LL, Temtamy S, Romeo G, Kate LP, et al. Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. *Genet Med*. 2011;**13**(9):841-7. doi: [10.1097/GIM.0b013e318217477f](https://doi.org/10.1097/GIM.0b013e318217477f). [PubMed: [2155946](https://pubmed.ncbi.nlm.nih.gov/2155946/)].
29. Daher S, El-Khairy L. Association of cerebral palsy with consanguineous parents and other risk factors in a Palestinian population. *East Mediterr Health J*. 2014;**20**(7):459-68. doi: [10.26719/2014.20.7.459](https://doi.org/10.26719/2014.20.7.459). [PubMed: [25023773](https://pubmed.ncbi.nlm.nih.gov/25023773/)].
30. Abu-Rabia S, Maroun L. The effect of consanguineous marriage on reading disability in the Arab community. *Dyslexia*. 2005;**11**(1):1-21. doi: [10.1002/dys.271](https://doi.org/10.1002/dys.271). [PubMed: [15747804](https://pubmed.ncbi.nlm.nih.gov/15747804/)].
31. Othman H, Saadat M. Prevalence of consanguineous marriages in Syria. *J Biosoc Sci*. 2009;**41**(5):685-92. doi: [10.1017/S0021932009003411](https://doi.org/10.1017/S0021932009003411). [PubMed: [19433003](https://pubmed.ncbi.nlm.nih.gov/19433003/)].
32. Motsinger AA, Ritchie MD. Multifactor dimensionality reduction: an analysis strategy for modelling and detecting gene-gene interactions in human genetics and pharmacogenomics studies. *Hum Genomics*. 2006;**2**(5):318-28. [PubMed: [16595076](https://pubmed.ncbi.nlm.nih.gov/16595076/)]. [PubMed Central: [PMC3500181](https://pubmed.ncbi.nlm.nih.gov/PMC3500181/)].
33. Lee S, Kim Y, Kwon MS, Park T. A Comparative Study on Multifactor Dimensionality Reduction Methods for Detecting Gene-Gene Interactions with the Survival Phenotype. *Biomed Res Int*. 2015;**2015**:671859. doi: [10.1155/2015/671859](https://doi.org/10.1155/2015/671859). [PubMed: [26339630](https://pubmed.ncbi.nlm.nih.gov/26339630/)]. [PubMed Central: [PMC4538337](https://pubmed.ncbi.nlm.nih.gov/PMC4538337/)].