



Predictive Factors of Infant Mortality Using Data Mining in Iran

Mahmoud Hajjipour ¹, Niloufar Taherpour ², Haleh Fateh ³, Ebrahim Yousefi ⁴, Koorosh Etemad ⁵, Fatemeh Zolfizadeh ⁶, Abdolhalim Rajabi ⁷, Tannaz Valadbeigi ⁸ and Yadollah Mehrabi ^{5, *}

¹Pediatric Gastroenterology, Hepatology and Nutrition Research Center, Research Institute for Children's Health, Shahid Beheshti University of Medical sciences, Tehran, Iran

²Prevention of Cardiovascular Disease Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

³Faculty of Computer Engineering, K.N. Toosi University of Technology, Tehran, Iran

⁴School of Applied Medical Science, Department of Health Information Technology and Management, Shahid Beheshti University of Medical Science, Tehran, Iran

⁵Department of Epidemiology, School of Public Health and safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁶Health Care Management, Mother and Child Welfare Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran

⁷Department of Epidemiology, Faculty of Health, Iran university of Medical Sciences, Tehran, Iran

⁸Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

* Corresponding author: Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: mehrabi@sbmu.ac.ir

Received 2020 August 16; Revised 2021 January 06; Accepted 2021 January 12.

Abstract

Objectives: Reducing infant mortality in the whole world is one of the millennium development goals. The aim of this study was to determine the factors related to infant mortality using data mining algorithms.

Methods: This population-based case-control study was conducted in eight provinces of Iran. A sum of 2,386 mothers (1,076 cases and 1,310 controls) enrolled in this study. Data were extracted from health records of mothers and filled with checklists in health centers. We employed several data mining algorithms such as AdaBoost classifier, Support Vector Machine, Artificial Neural Networks, Random Forests, K-nearest neighborhood, and Naïve Bayes in order to recognize the important predictors of infant death; binary logistic regression model was used to clarify the role of each selected predictor.

Results: In this study, 58.7% of infant mortalities occurred in rural areas, that 55.6% of them were boys. Moreover, Naïve Bayes and Random Forest were highly capable of predicting related factors among data mining models. Also, the results showed that events during pregnancy such as dental disorders, high blood pressure, loss of parents, factors related to infants such as low birth weight, and factors related to mothers like consanguineous marriage and gap of pregnancy (< 3 years) were all risk factors while the age of pregnancy (18 - 35 year) and a high degree of education were protective factors.

Conclusions: Infant mortality is the consequence of a variety of factors, including factors related to infants themselves and their mothers and events during pregnancy. Owing to the high accuracy and ability of modern modeling compared to traditional modeling, it is recommended to use machine learning tools for indicating risk factors of infant mortality.

Keywords: Infant Mortality, Risk Factors, Machine Learning, Logistic Regression Model

1. Background

Infant mortality is one of the most important health indicators that affect the level of health and development of countries (1). Infant mortality is a death that occurs before the age of one, and the infant mortality rate is the number of deaths among infants that occur in every 1,000 live births (2). Reducing child mortality, especially infant mortality in the whole world, is one of the millennium development goals (3). According to the World Health Organization (WHO) reports in 2017, 4.1 million (75% of all under-five mortality) occurred within the first year of life. African region had the highest risk of child death under

the age of one (51 per 1,000 live births), while the European region showed eight deaths per 1,000 live births. Thus child death in the African region was six times higher than European region (4). According to the reports by CDC in 2016, infant mortality rates per 1,000 live births by race and ethnicity were: 11.4 None-Hispanic black, 9.4 American-Indian/Alaska native, 7.4 native Hawaiian or other Pacific Islander, 4.9 None-Hispanic whites, and 3.6 Asians (2). Globally, in 2017, the infant mortality rate decreased from 65 in 1990 to 29 deaths per 1,000 live births; in the United States, the infant mortality rate was 5.9 per 1000 live births in 2016 (2, 4). In Iran, the infant mortality rate has been decreased from 44 per 1,000 live births in 1990 to 13 in 2017

(5), with the highest rate in neonates (6). According to the report by WHO in 2008, the main cause of death in infants were prenatal causes, diarrheal diseases, lower respiratory infections, Malaria, congenital anomalies, Pertussis, HIV/AIDS, Tetanus, Meningitis, Measles, protein-energy malnutrition, Syphilis, endocrine disorders, Tuberculosis, and upper respiratory infections (7). Preterm birth, low birth weight, sudden infant death syndrome, maternal pregnancy complications, injuries and suffocation are among the causes of infant mortality reported by CDC (2). It is trivial that all of these causes may be changed due to demographic transition, socio-economic, political, and health status of countries during the time (8).

In this study, we used data mining techniques to extract hidden information from a large volume of data and to predict risk factors related to infant mortality (9). Machine learning or data mining is defined as enabling computers to produce evidence for successful decision-making based on past experiences. In recent years, a rapid increase in storage capacity and processing power helped machine learning to exhibit an impressive development (10). Advances in data analysis methods contributed to machine learning enables computers to make an accurate and reliable prediction using even a very small sample size (11). Machine learning techniques such as Decision Trees (DT), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Naïve Bayes, and Artificial Neural Network (ANN) have been used for classification and prediction about detection, treatment, morbidity, and mortality among patients or population (12).

2. Objectives

As infant mortality is preventable and most of the risk factors are controllable by health policymaking in the community, we performed a study to identify the risk factors and determinants of infant mortality using data mining techniques for prediction, prevention, and promotion of mothers and especially children's health.

3. Methods

This population-based case-control study was conducted in eight provinces of Iran, including Fars, Golestan, Hormozgan, Kohgiluyeh, and Boyer-Ahmad, South Khorasan Kermanshah, Hamedan, and Yazd in 2017. In a stratified cluster random sampling, eight out of 31 provinces from North, East, West, South, and center of Iran were selected. Then five districts were chosen again from North,

East, West, South, and Center of each selected province. Thereafter, we randomly selected one urban and one rural health center from each sampled district. According to literature review, considering mother's education < 5 years as a risk factor ($P_0 = 0.58$, $P_1 = 0.68$, $Z_{0.95} = 1.96$, $Z_{(1-\beta)} = 0.90$, design effect = 2) and using sample size formula for comparison of proportions, the sample size was estimated as 508 for each of case and control groups. By taking design effect = 2 into account, the sample size for each group reached 1,016. We added some more samples to replace the probable incomplete data. The final sample size was 2,386 (1,076 cases and 1,310 controls). The case group included mothers who had lost an under one-year-old child due to disease or congenital disorder. The control group included mothers who had at least a live child with the age of less than one and had not experienced infant death. Subjects in case and control groups were randomly selected from mothers who were referred to the chosen health centers. Data were collected using researcher-made checklists from mother's health records. Participants with poor reading or writing skills were interviewed face to face by a trained interviewer to complete the checklist. Checklists included data of demographic status and diseases or disorders as well as the history of mothers and their children.

3.1. Statistical Analysis

Data accuracy and disparities were evaluated. In this study, the outcome variable was considered vital status of infant, while predictor variables included demographic status of mother, history of diseases, events during pregnancy, violence or trauma on mother, and medical records of mothers and their infants such as sex, birth weight, and so on. We employed several data mining algorithms such as AdaBoost classifier, Support Vector Machine (SVM), Artificial Neural Networks (ANN), Random Forests (RF), K-Nearest Neighborhood, and Naïve Bayes to recognize the important predictors of infant death. Sensitivity, specificity, classification accuracy, and F1 score with the following formulas were used for measuring and comparing the performance of the classification methods.

$$\text{Sensitivity (Recall)} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FP + FN} \quad (4)$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \quad (5)$$

Here TP, TN, FP, and FN stands for true positive, true negative, false positive, and false negative values, respectively (13). Afterward, we used six data mining algorithms to indicate important variables related to infant mortality. Thus factors like mother age at pregnancy, living place, mother literacy and job, consanguineous marriage, gap of pregnancy, worst life event, smoking during pregnancy, child sex, twins, dental disorders, psychological syndrome, gestational diabetes, high blood pressure, and anemia during pregnancy were selected as predictor factors. Among them, 16 variables gained acceptable important scores according to the results of data mining models.

After recognition of important predictors, the binary logistic regression model was used to clarify the role of each selected predictor. Area under receiver operating characteristic curve (ROCAUC) and Hosmer-Lemeshow test were used to measure the calibration and discrimination ability of models.

Data analyses were carried out using STATA-Version 14 and Python 2.7. IDLE software, and the significance level was considered $P < 0.05$.

4. Results

Of 2,386 studied mothers (1,076 cases and 1,310 controls), 58.7% of the cases and 54.9% of the controls were living in rural areas.

In the case group, 14.2% of mothers were over 35 years old at the time of pregnancy compared to 8.8% in the control group. A total of 52.4% of dead infants were boys. Consanguineous marriage among subjects who had experienced infant death (33.6%) was more prevalent than subjects who had no such experience (25.6%). A total of 4.82% of mothers reported smoking during their pregnancy, and a total of 25.4% of them reported a history of diseases such as anemia, and 5.4% of them had diabetes during their pregnancy. The characteristics of the study subjects in case and control groups are shown in Table 1. After running six data mining algorithms, variables that gained an acceptable importance score were identified.

For better comparison, we used seven statistical models (logistic regression and six data mining models) for the prediction of outcome using all 16 important predictor variables. The results of model effectiveness shown in Table 2 represent that Naïve Bayes and Random Forest had the highest AUC, F1-score, precision, and sensitivity

among data mining algorithms, respectively. These algorithms had high power in predictions of related factors among data mining models (Table 2). The impact value and Odds ratio of each variable using univariate and multivariate regression models are represented in Table 3. Logistic regression model showed the adjusted Odds ratio of consanguineous marriage as 1.44 (95% CI: 1.18-1.76), short pregnancy gap or first pregnancy, worst life event as 1.65 (95% CI: 1.36-2.00), dental disorders as 2.49 (95% CI: 1.90-3.27) and high blood pressure during pregnancy as 1.62 (95% CI: 1.11-2.38) and neonatal weight of under 2500g as 8.13 (95% CI: 6.34-10.42) significantly increased the risk of infant mortality. On the other hand, maternal age of 18-30 and 30-35 years old and mother's literacy (high school and university degree) all acted as protective factors for infant mortality (Table 3).

5. Discussion

In our study, we indicated important variables that were related to infant mortality using data mining algorithms and a logistic regression model. Based on the results, Naïve Bayes among data mining algorithms had better performance in terms of precision, AUC, F1-score, and sensitivity compared to other algorithms; also, the results of logistic regression were similar to data mining algorithms. So we can say, even in studies with a large sample size, traditional models (logistic regression) are similar to modern models (data mining), which have high potency to release accurate and fit models to predict important related factors. In similar studies, the results of comparing modern and traditional modeling showed that data mining methods did not have any advantage over logistic regression for prediction. The results of logistic regression and data mining (value of AUC and Precision) were close together, but in some cases, logistic regression had a better performance (14, 15); however, some articles reported that data mining models (Naïve Bayes Network and Artificial Neural Network) were more accurate and efficient compared to logistic regression model (16, 17). In the present study, the results of data mining models showed that important factors related to infant mortality were mother's age of pregnancy, place of living, mother's literacy, mother's job, consanguineous marriage, gap of pregnancy, worst life event, smoking during pregnancy, sex of child, twins, dental disorders, psychological syndrome, high blood pressure during pregnancy, gestational diabetes, and anemia during pregnancy, respectively. Therefore, infant mortality among mothers with normal age

Table 1. Demographic Characteristics of Subjects in Case and Control Groups^a

Variables	Cases (N = 1,076)	Controls (N = 1,310)
Child		
Infant sex (girl)	478 (44.4)	615 (46.9)
Birth weight		
Normal (2,500 - 4,000 g)	599 (55.7)	1169 (89.2)
Low birth weight (< 2,500 g)	463 (43.0)	122 (9.3)
High birth weight (> 4,000g)	14 (1.3)	19 (1.5)
Twins or more	86 (8.0)	45 (3.4)
Mother		
Living place (rural)	632 (58.7)	719 (54.9)
Age of getting pregnant		
< 18	30 (2.80)	28 (2.1)
18 - 30	681 (63.3)	859 (65.6)
30 - 35	212 (19.7)	308 (23.5)
> 35	153 (14.2)	115 (8.8)
Job		
Clerk	39 (3.6)	108 (8.2)
Housewife	1008 (93.7)	1174 (89.6)
Labor	29 (2.7)	28 (2.1)
Literacy		
Illiterate	71 (6.6)	53 (4.0)
Elementary school	277 (25.7)	252 (19.2)
Middle school	271 (25.2)	272 (20.8)
High school	369 (34.3)	551 (42.1)
University degree	88 (8.2)	182 (13.9)
Consanguineous marriage	362 (33.6)	336 (25.6)
Pregnancy gap (y)		
> 3	305 (28.3)	467 (35.6)
1-3	305 (28.3)	357 (27.3)
< 1	78 (7.2)	46 (3.5)
First pregnancy	388 (36.1)	440 (33.6)
Events During Pregnancy		
Smoking	60 (5.6)	55 (4.2)
Worst life event		
Experience of family loss	439 (40.8)	413 (31.5)
High blood pressure	103 (9.6)	62 (4.7)
Diabetes	58 (5.4)	70 (5.3)
Anemia	255 (23.7)	350 (26.7)
Dental disorders	222 (20.6)	130 (9.9)
Psychological syndromes	26 (2.4)	15 (1.1)

^aValues are expressed as No. (%).

Table 2. F1-Score, Precision, Sensitivity, and Area Under ROC Curve (AUC) of Logistic Regression and Six Data Mining Algorithms

Classifier	AUC	F1-Score	Precision	Sensitivity (Recall)
Adaboost classifier	0.749	0.696	0.696	0.697
Support vector machine	0.699	0.633	0.642	0.632
Neural networks	0.754	0.697	0.696	0.697
Random forests	0.777	0.721	0.723	0.723
K-nearest neighbor	0.738	0.682	0.692	0.689
Naive bayes	0.785	0.720	0.728	0.725
Logistic regression	0.788	0.722	0.727	0.725

during pregnancy (18 - 35 years) was 40% lower than the mothers with age 35 years old and over. Pregnancy in < 18 years old due to biological and psychological insufficiency of mothers and chance of low birth weight and pregnancy in > 35 years due to high probability of born with cognitive disorders can increase the stillbirth and infant mortality; so normal range of mother's age in pregnancy can be a protective factor for infant mortality (18, 19). Another important factor was consanguineous marriage; in other words, consanguineous marriage among parents increased the chance of infant mortality to 44% compared to non-consanguineous marriage. Similar studies reported that consanguineous marriage is responsible for congenital disorders and genetic diseases like Down syndrome, thalassemia, asthma, mental disorders, heart diseases, gastrointestinal disorders, and hearing deficiency that influence the health and survival of infants and children (20, 21). Gap of pregnancy was significantly associated with infant mortality so that short intervals between pregnancies (< 3 years) can increase the risk of infant mortality. In other words, inadequate maternal recovery time and its complication like anemia, adverse psychological effect of delivery, inadequate mother care for infants, cessation of breastfeeding, and spreading infectious diseases among individuals are consequences of short birth interval which affect infant mortality (22). Also, in this study, we found that first pregnancy increased the chance of infant mortality to 53% that can be due to low maternal experience in infant care. Mother education can be a protective factor for infant mortality. Because university education and high school education reduced infant mortality to 60% and 56%, respectively; education can increase connections of mothers with resources for infant health, awareness of healthy behaviors, and access to health services (23). History of the worst life events like loss of parents during pregnancy among mothers was significantly associated with infant mortality, and 65% increased the chance of infant mor-

tality. Psychological traumatic events such as loss of parents affect physical and mental health, loneliness, and infant poor care, which all can affect infant mortality (24, 25). The results showed that mortality among infants with low birth weight (< 2500 g) was 8.13 more than the infants with normal birth weight (2500 - 4000 g); it is due to the vulnerability of infants to various diseases and death (26). Finally, a significant relation was found between a history of diseases during pregnancy like dental disorders and high blood pressure with infant mortality. As a result, infant mortality among mothers with a history of dental disorders and with high blood pressure during pregnancy were 2.49 and 1.62, respectively, more than mothers without this complication. Similar studies reported that periodontal disease and low oral health can indirectly influence low birthweight; thus, dental disorders like periodontal disease in pregnant women with reservoir of microorganisms can be a risk factor for adverse outcomes like low birth weight and, finally, neonatal and infant mortality (27, 28). Also, high blood pressure during pregnancy (systolic blood pressure \geq 140 mmHg and/or diastolic blood pressure \geq 90 mmHg) affects a mother's health and her infant. High blood pressure is responsible for preeclampsia, stroke among mothers, and preterm delivery that lead to infant mortality or stillbirth (29).

Finally, variables like infant's sex, twins, smoking during pregnancy, gestational diabetes, gestational anemia, living place (urban or rural), mother's job, and psychological symptoms during pregnancy were all indicated as important factors related to infant mortality, but in the logistic regression model, a significant relation was not found considering these factors. Missing data due to incomplete checklist and response bias and underestimation due to unwillingness of participants to report factors with strong social stigma among women such as smoking were the limitation of the study.

Table 3. Crude and Adjusted Odd Ratio for Infant Mortality: Logistic Regression Model

Variables	Crude OR (95% CI)	P - Value	Adjusted OR (95% CI)	P - Value
Living place (rural)	1.17 (0.99 - 1.37)	0.05	1.18 (0.96 - 1.44)	0.10
Maternal age				
< 18	0.80 (0.45 - 1.42)	0.45	0.63 (0.32 - 1.24)	0.18
18 - 30	0.59 (0.45 - 0.77)	< 0.001	0.60 (0.43 - 0.84)	0.003
30 - 35	0.51 (0.38 - 0.69)	< 0.001	0.60 (0.43 - 0.85)	0.005
> 35	Ref	-	Ref	-
Mother's job				
Clerk	0.42 (0.28 - 0.61)	< 0.001	0.67 (0.42 - 1.06)	0.09
Labor	1.20 (0.71 - 2.04)	0.48	1.25 (0.68 - 2.29)	0.46
Housewife	Ref	-	Ref	-
Consanguineous marriage	1.46 (1.23 - 1.75)	< 0.001	1.44 (1.18 - 1.76)	< 0.001
Gap of pregnancy (y)				
1 - 3	1.30 (1.06 - 1.61)	0.01	1.42 (1.11 - 1.82)	0.005
< 1	2.56 (1.75 - 3.84)	< 0.001	2.70 (1.72 - 4.23)	< 0.001
First pregnancy	1.35 (1.10 - 1.64)	0.003	1.53 (1.19 - 1.98)	0.001
> 3	Ref	-	Ref	-
Child sex (girl)	0.90 (0.76 - 1.06)	0.21	0.92 (0.77 - 1.11)	0.42
Smoking during pregnancy (yes)	1.34 (0.92 - 1.96)	0.11	1.11 (0.72 - 1.72)	0.61
Mother education				
Elementary school	0.82 (0.55 - 1.21)	0.32	0.89 (0.56 - 1.39)	0.60
Middle school	0.74 (0.50 - 1.10)	0.14	0.71 (0.45 - 1.13)	0.15
High school	0.49 (0.34 - 0.73)	< 0.001	0.44 (0.28 - 0.69)	< 0.001
University degree	0.36 (0.23 - 0.55)	< 0.001	0.40 (0.23 - 0.69)	0.001
Worst life event experienced during pregnancy				
Loss of family member	1.49 (1.26 - 1.77)	< 0.001	1.65 (1.36 - 2.00)	< 0.001
Birth weight				
Low birth weight (< 2500 g)	7.40 (5.92 - 9.25)	< 0.001	8.13 (6.34 - 10.42)	< 0.001
High birth weight (> 4000 g)	1.43 (0.71 - 2.88)	0.30	1.28 (0.60 - 2.69)	0.51
Normal (2500 - 4000 g)	Ref	-	Ref	-
Twins or more	2.44 (1.68 - 3.53)	< 0.001	0.84 (0.54 - 1.31)	0.45
Psychological symptoms during pregnancy	2.13 (1.12 - 4.05)	0.02	1.42 (0.67 - 3.00)	0.35
Diabetes during Pregnancy	1.18 (0.82 - 1.71)	0.35	1.19 (0.77 - 1.83)	0.41
Anemia during pregnancy	0.85 (0.70 - 1.02)	0.09	0.90 (0.73 - 1.12)	0.37
Dental disorders during pregnancy	2.35 (1.86 - 2.98)	< 0.001	2.49 (1.90 - 3.27)	< 0.001
High blood pressure during pregnancy	2.13 (1.53 - 2.95)	< 0.001	1.62 (1.11 - 2.38)	0.01

5.1. Conclusions

Infant mortality is a multifactorial outcome, which includes infants, mothers, and events during pregnancy. Events during pregnancy such as dental disorders, high

blood pressure, and loss of parents, factors related to infants such as low birth weight (< 2500 g), factors related to mothers like consanguineous marriage, and gap of pregnancy (< 3 years) were risk factors, while the age of preg-

nancy (18 - 35 year) and a high degree of education were protective factors. Thus due to the high accuracy and potency of modern modeling similar to traditional modeling (logistic regression models), we can use machine learning for indicating related factors to infant mortality; consequently, these factors can be prevented.

Acknowledgments

The authors would like to thank the medical sciences universities and health centers of eight provinces who participated in this project and provided their services with dedication.

Footnotes

Authors' Contribution: Hajipour: Data analysis, data acquisition, the focal point of data gathering in South-Khorasan and Hamedan; Taherpour: Drafting the manuscripts in Persian and revising the manuscripts in English, preparing results and tables; Fateh and Ebrahimi: Data analysis (data mining); Etemad: Data acquisition and focal point of data gathering in Golestan and Kohgiluyeh and Boyer-Ahmad provinces; Zolfizadeh: Data acquisition, quality control of data in Hormozgan province; Rajabi: Data acquisition, the focal point of data gathering in Yazd and Kermanshah provinces; Valadbeigi: Data acquisition, focal point of data gathering in Fars province; Mehrabi: Supervisor, designing the project, and revising the manuscripts in English.

Conflict of Interests: The authors declare that they have no competing interests.

Ethical Approval: Ethical approval was granted by the National Institutes for Medical Research Development (NIMAD), Tehran, Iran (IR.NIMAD.REC.1398.033).

Funding/Support: This research was supported by Elite Researcher Grant Committee under award number 958930 from the National Institutes for Medical Research Development (NIMAD), Tehran, Iran.

Informed Consent: All participants were provided written informed consent before the interview; this included permission to use anonymized quotations in publications.

References

- Zewdie SA, Adjiwanou V. Multilevel analysis of infant mortality and its risk factors in South Africa. *Int J Popul Stud.* 2017;3(2):43. doi: [10.18063/ijps.v3i2.330](https://doi.org/10.18063/ijps.v3i2.330).
- CDC. *Infant Mortality*. Centers for Disease Control and Prevention; 2018, [cited 2020 Nov]. Available from: <https://www.cdc.gov/reproductivehealth/MaternalInfantHealth/InfantMortality.htm>.
- WHO. *Millennium development goals (MDGs)*. World Health Organization; 2018, [cited 2020 Nov]. Available from: http://www.who.int/topics/millennium_development_goals/about/en/.
- WHO. *Infant mortality (situation and trend)*. World Health Organization; 2018, [cited 2020 Nov]. Available from: http://www.who.int/gho/child_health/mortality/neonatal_infant_text/en/.
- UNICEF. *Level and trends in child mortality*. New York: United Nations Children's Fund; 2018. 33 p.
- Ansari-Moghaddam A, Sadeghi-Bojd S, Imani M, Movahedinia S, Pour-rashidi A, Mohammadi M. A multivariate analysis of factors associated with infant mortality in South-East of Iran. *J Pak Med Assoc.* 2014;64(10):1123-6. [PubMed: [25823149](https://pubmed.ncbi.nlm.nih.gov/25823149/)].
- WHO. Peden M, Oyegbite K, Ozanne-Smith J, Hyder AA, Branche C, Rahman A, et al., editors. *World report on child injury prevention*. Geneva: World Health Organization; 2008.
- Montgomery MR, Cohen B; National Research Council (US) Committee on Population. *The Impact of Infant and Child Mortality Risk on Fertility. From Death to Birth: Mortality Decline and Reproductive Change*. Washington (DC): National Academies Press (US); 1998.
- Sahle G. Ethiopian maternal care data mining: discovering the factors that affect postnatal care visit in Ethiopia. *Health Inf Sci Syst.* 2016;4:4. doi: [10.1186/s13755-016-0017-2](https://doi.org/10.1186/s13755-016-0017-2). [PubMed: [27217953](https://pubmed.ncbi.nlm.nih.gov/27217953/)]. [PubMed Central: [PMC4876568](https://pubmed.ncbi.nlm.nih.gov/PMC4876568/)].
- Bastanlar Y, Ozuysal M. Introduction to machine learning. *Methods Mol Biol.* 2014;1107:105-28. doi: [10.1007/978-1-62703-748-8_7](https://doi.org/10.1007/978-1-62703-748-8_7). [PubMed: [24272434](https://pubmed.ncbi.nlm.nih.gov/24272434/)].
- Li W, Huang Y, Zhuang BW, Liu GJ, Hu HT, Li X, et al. Multiparametric ultrasomics of significant liver fibrosis: A machine learning-based analysis. *Eur Radiol.* 2019;29(3):1496-506. doi: [10.1007/s00330-018-5680-z](https://doi.org/10.1007/s00330-018-5680-z). [PubMed: [30178143](https://pubmed.ncbi.nlm.nih.gov/30178143/)]. [PubMed Central: [PMC6510867](https://pubmed.ncbi.nlm.nih.gov/PMC6510867/)].
- Liou DM, Chang WP. Applying data mining for the analysis of breast cancer data. *Methods Mol Biol.* 2015;1246:175-89. doi: [10.1007/978-1-4939-1985-7_12](https://doi.org/10.1007/978-1-4939-1985-7_12). [PubMed: [25417087](https://pubmed.ncbi.nlm.nih.gov/25417087/)].
- LeDell E, van der Laan MJ, Petersen M. AUC-maximizing ensembles through metalearning. *Int J Biostat.* 2016;12(1):203-18. doi: [10.1515/ijb-2015-0035](https://doi.org/10.1515/ijb-2015-0035). [PubMed: [27227721](https://pubmed.ncbi.nlm.nih.gov/27227721/)]. [PubMed Central: [PMC4912128](https://pubmed.ncbi.nlm.nih.gov/PMC4912128/)].
- Kuhle S, Maguire B, Zhang H, Hamilton D, Allen AC, Joseph KS, et al. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. *BMC Pregnancy Childbirth.* 2018;18(1):333. doi: [10.1186/s12884-018-1971-2](https://doi.org/10.1186/s12884-018-1971-2). [PubMed: [30111303](https://pubmed.ncbi.nlm.nih.gov/30111303/)]. [PubMed Central: [PMC6094446](https://pubmed.ncbi.nlm.nih.gov/PMC6094446/)].
- Witteveen A, Nane GF, Vliegen IMH, Siesling S, I. Jzerman MJ. Comparison of logistic regression and bayesian networks for risk prediction of breast cancer recurrence. *Med Decis Making.* 2018;38(7):822-33. doi: [10.1177/0272989X18790963](https://doi.org/10.1177/0272989X18790963). [PubMed: [30132386](https://pubmed.ncbi.nlm.nih.gov/30132386/)].
- Langarizadeh M, Moghbeli F. Applying naive bayesian networks to disease prediction: A systematic review. *Acta Inform Med.* 2016;24(5):364-9. doi: [10.5455/aim.2016.24.364-369](https://doi.org/10.5455/aim.2016.24.364-369). [PubMed: [28077895](https://pubmed.ncbi.nlm.nih.gov/28077895/)]. [PubMed Central: [PMC5203736](https://pubmed.ncbi.nlm.nih.gov/PMC5203736/)].
- Li G, Zhou X, Liu J, Chen Y, Zhang H, Chen Y, et al. Comparison of three data mining models for prediction of advanced schistosomiasis prognosis in the Hubei province. *PLoS Negl Trop Dis.* 2018;12(2). e0006262. doi: [10.1371/journal.pntd.0006262](https://doi.org/10.1371/journal.pntd.0006262). [PubMed: [29447165](https://pubmed.ncbi.nlm.nih.gov/29447165/)]. [PubMed Central: [PMC5831639](https://pubmed.ncbi.nlm.nih.gov/PMC5831639/)].
- Neal S, Channon AA, Chintsanya J. The impact of young maternal age at birth on neonatal mortality: Evidence from 45 low and middle income countries. *PLoS One.* 2018;13(5). e0195731. doi: [10.1371/journal.pone.0195731](https://doi.org/10.1371/journal.pone.0195731). [PubMed: [29791441](https://pubmed.ncbi.nlm.nih.gov/29791441/)]. [PubMed Central: [PMC5965834](https://pubmed.ncbi.nlm.nih.gov/PMC5965834/)].

19. Finlay JE, Ozaltin E, Canning D. The association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemia for first births: evidence from 55 low- and middle-income countries. *BMJ Open*. 2011;**1**(2). e000226. doi: [10.1136/bmjopen-2011-000226](https://doi.org/10.1136/bmjopen-2011-000226). [PubMed: [22021886](https://pubmed.ncbi.nlm.nih.gov/22021886/)]. [PubMed Central: [PMC3191600](https://pubmed.ncbi.nlm.nih.gov/PMC3191600/)].
20. Bener A, Hussain R, Teebi AS. Consanguineous marriages and their effects on common adult diseases: studies from an endogamous population. *Med Princ Pract*. 2007;**16**(4):262–7. doi: [10.1159/000102147](https://doi.org/10.1159/000102147). [PubMed: [17541290](https://pubmed.ncbi.nlm.nih.gov/17541290/)].
21. Koc I, Eryurt MA. The causal relationship between consanguineous marriages and infant mortality in Turkey. *J Biosoc Sci*. 2017;**49**(4):536–55. doi: [10.1017/S002193201600033X](https://doi.org/10.1017/S002193201600033X). [PubMed: [27382977](https://pubmed.ncbi.nlm.nih.gov/27382977/)].
22. van Soest A, Saha UR. Relationships between infant mortality, birth spacing and fertility in Matlab, Bangladesh. *PLoS One*. 2018;**13**(4). e0195940. doi: [10.1371/journal.pone.0195940](https://doi.org/10.1371/journal.pone.0195940). [PubMed: [29702692](https://pubmed.ncbi.nlm.nih.gov/29702692/)]. [PubMed Central: [PMC5922575](https://pubmed.ncbi.nlm.nih.gov/PMC5922575/)].
23. Khajavi A, Pishgar F, Dehghani M, Naderimagham S. Socioeconomic inequalities in neonatal and postneonatal mortality: Evidence from rural Iran, 1998–2013. *Int J Equity Health*. 2017;**16**(1):83. doi: [10.1186/s12939-017-0570-6](https://doi.org/10.1186/s12939-017-0570-6). [PubMed: [28521753](https://pubmed.ncbi.nlm.nih.gov/28521753/)]. [PubMed Central: [PMC5437686](https://pubmed.ncbi.nlm.nih.gov/PMC5437686/)].
24. Voland E, Beise J. Opposite effects of maternal and paternal grandmothers on infant survival in historical Krummhörn. *Behav Ecol Sociobiol*. 2002;**52**(6):435–43. doi: [10.1007/s00265-002-0539-2](https://doi.org/10.1007/s00265-002-0539-2).
25. Hajipour M, Taherpour N, Payandeh A, Etemad K, Rezaeian S, Zolfizadeh F, et al. Injury- and non-injury-related factors associated with under five mortality in Iran: A multilevel study. *Health Scope*. 2020;**9**(2). doi: [10.5812/jhealthscope.99949](https://doi.org/10.5812/jhealthscope.99949).
26. Watkins WJ, Kotecha SJ, Kotecha S. All-cause mortality of low birth-weight infants in infancy, childhood, and adolescence: Population study of England and Wales. *PLoS Med*. 2016;**13**(5). e1002018. doi: [10.1371/journal.pmed.1002018](https://doi.org/10.1371/journal.pmed.1002018). [PubMed: [27163787](https://pubmed.ncbi.nlm.nih.gov/27163787/)]. [PubMed Central: [PMC4862683](https://pubmed.ncbi.nlm.nih.gov/PMC4862683/)].
27. Meqa K, Dragidella F, Disha M, Sllamniku-Dalipi Z. The Association between Periodontal Disease and Preterm Low Birthweight in Kosovo. *Acta Stomatol Croat*. 2017;**51**(1):33–40. doi: [10.15644/asc51/1/4](https://doi.org/10.15644/asc51/1/4). [PubMed: [28740268](https://pubmed.ncbi.nlm.nih.gov/28740268/)]. [PubMed Central: [PMC5506254](https://pubmed.ncbi.nlm.nih.gov/PMC5506254/)].
28. Haerian-Ardakani A, Eslami Z, Rashidi-Meibodi F, Haerian A, Dallalnejad P, Shekari M, et al. Relationship between maternal periodontal disease and low birth weight babies. *Iran J Reprod Med*. 2013;**11**(8):625–30. [PubMed: [24639799](https://pubmed.ncbi.nlm.nih.gov/24639799/)]. [PubMed Central: [PMC3941365](https://pubmed.ncbi.nlm.nih.gov/PMC3941365/)].
29. Braunthal S, Brateanu A. Hypertension in pregnancy: Pathophysiology and treatment. *SAGE Open Med*. 2019;**7**:2050312119843700. doi: [10.1177/2050312119843700](https://doi.org/10.1177/2050312119843700). [PubMed: [31007914](https://pubmed.ncbi.nlm.nih.gov/31007914/)]. [PubMed Central: [PMC6458675](https://pubmed.ncbi.nlm.nih.gov/PMC6458675/)].