RISK OF INFANT MORTALITY ANALYSIS BY FUZZY LOGIC APPROACH AND PERFORMANCE STATUS OF THE MODEL

Debajyoti Bora[1] and Manash Pratim Barman[2,*]

[1]Department of Statistics, PDUAM, Eraligool, Karimganj,

[2]Department of Statistics, Dibrugarh University, Dibrugarh,

**Corresponding Author*: barman_manash@rediffmail.com

Abstract

Infant mortality rate is considered one of the leading indicators for assessment of population health, especially with regard to quality of care provided to pregnant women and newborns. Objective of this research is to study possible risk of infant mortality through fuzzy approach. In this context, a predictive model is developed here to estimate risk of infant mortality using fuzzy logic approach on basis of fuzziness of three variables viz. Birth weight, Mother's age and Mother's Anemia. Mamdani method is used for inference and Centroid method is used for defuzzification of the risk of infant mortality given by the model. Validation of the model is done in three stages; ROC curve, statistical test and individual observation by taking real data from NFHS-III. Model estimates low risk to the infants who survived than those who had died. Accuracy of the model is 0.934 (95 per cent CI: 0.84-1.00) ($p < 0.001$). Accuracy is higher with risk below 32 per cent, corresponding to 0.91 (95 per cent CI: 0.85-0.95) in respect to sensitivity, 0.71 (95 per cent CI: 0.35-0.92) in respect to specificity, 0.99 positive predictive value and 0.26 negative predictive value. Output of the fitted model determines its applicability in predicting demographic events.

Key words: Infant Mortality, Risk Factor, Fuzzy Inference, Linguistic Variable.

1. INTRODUCTION

The mortality during first year of life of a live born is considered as one of the most sensitive indicators of socio-economic development and the general health conditions of a community or a country. Measure of infant mortality also reflects the living standard of the people and effectiveness of government health services for improving maternal and child health in a country. The public health specialists considered infant mortality as of greater importance than other mortalities such maternal mortality, under-five mortality etc., because it is the single, largest category of mortality. The changes in specific health interventions affect infant mortality more rapidly and directly, as a result it may change more dramatically than the crude death in the population. Therefore, an argent need is here for developing models through the involvement of existing factors of infant mortality. The models not only explain the relationship between possible associated factors but it also provides estimates of possible risks by combination of the factors. By identifying the risk of death associated with the newborn infant processing different demographic and health characteristics may facilitate academicians, administrators, policy makers and programme planners to implement initiatives for improvement of the status of the infants.

Logistic regression model has been applied using dichotomous independent variables, such as present or absent, yes or no to estimate the risk of infant mortality (Silva & Mathias, 2014, Ezeh et al., 2015, Phipps et al., 2002). In such method the classification of mortality factors do not show any significant differences in construction and analyzing the model. For example if we consider birth weight as a mortality factor of infant death, then birth weight 2,489g of an infant and birth weight 2,556g of another infant are classically classified as low birth weight and normal birth weight respectively, which does not show any significant differences. However, the best and most useful description of mortality factors often comprise linguistic terms that are inevitably vague, such as age (early or late) and birth weight (high or low). Recently, for this reason, fuzzy logic got the attention in epidemiology, public health and demographic due to the ability to deals with linguistic terms which are unavoidably vague and uncertain. In this approach each element can be classified into several categories, with different membership values and in the absence of precise and absolute information, fuzzy logic can be effectively used for modeling and predicting future events.

A fuzzy linguistic model is a rule base system which uses fuzzy theory and fuzzy logic to address the concern problem. It contains four main components;

- Fuzzifier: it translates crisp input value(s) into fuzzy value(s).
- Knowledge Base: it includes both fuzzy inference rule sets and their membership functions representing the fuzzy sets of linguistic variable. Fuzzy inference rules are usually expressed in the form 'IF-THEN' that defines the connection between input and output variable.
- Inference Engine: It applies a fuzzy reasoning to obtain a fuzzy output(s). The inference engine assesses all fuzzy rules in the rule base and combines the weighted consequents of all relevant rules into a single output fuzzy set.
- Defuzzifier: It translates the fuzzy output value(s) into crisp value(s).

For instance, fuzzy logic has been successfully applied in the diagnosis of various diseases and demographic event prediction using associated factors as input variables (Diab & Saade, 2005, Yilmaz & Ayan 2013, Nascimento et al., 2009, Chaves & Nascimento, 2014, Nascimento & Ortega, 2002, Bora & Barman, 2017). Thus here is an attempt to develop a fuzzy linguistic model for estimating risk of infant mortality using fuzzy logic approach based on birth weight, mother's age and mother's anaemia.

2. METHODOLOGY

Input and output variable for the model:

Form literature survey (Phipps et al., 2002, Maia et al., 2012, Mombelli et al., 2012, Barros & Victora, 2008, Kassar et al., 2013, Mathews & MacDorman, 2007, Cooper et al., 1995, Benjamin et al., 2009, Barman & Saikia, 2013, Sharma & Shankar, 2010) three factors are taken as inputs variables for the fuzzy model to estimate the risk of infant mortality, viz.- new born birth weight, age of the mother and Anaemia (Haemoglobin in g/dl) level of mother. To convert crisp input variables into fuzzy sets previous studies (Nascimento et al., 2009, Chaves & Nascimento, 2014, Nascimento & Ortega, 2002, Rajeswari et al., 2015, Smith, 2012, Kalaivani, 2009, National Institutes of Health, 2011) are followed.

Selected inputs are stratified for the construction of the predictive fuzzy model. The variable birth weight categorized into three fuzzy sets; which are very low birth weight when weights are below 1500g, low birth weight when weights are between 1500g and 2500g and normal birth weight when weights are greater than and equal to 2500g. The variable mother's age is divided into three fuzzy sets such as early (≤ 18 years), normal (between 18 years and 35years) and late (35years and above). The variable mother's anaemia is categorized into four fuzzy sets as severe if the level of anemia is lies between 4.4g/dl to 6.9g/dl, moderate if the level of anaemia is lies between 7.0g/dl to 10.0g/dl, mild if the level of anaemia is lies between 10.0g/dl to 11.0g/dl and not anaemic if the level of anaemia is more than 11.0g/dl. The output i.e. the risk of infant mortality is also categorized into three linguistic labels; which are low, middle and high.

Fuzzy Inference Processes:

Input and output variables are converted to fuzzy sets by calculating the membership degree. Fuzzy rules are formulated by considering a stratified analysis of NFHS-II data taking the studied variables (see Table 1).

Table 1: Fuzzy Rules of the model

				Anemia			
				Severe	Moderate	Mild	Not Anaemic
Age	Early	Birth Weight	Very Low	*High*	*High*	*High*	*High*
			Low	*High*	*High*	*High*	*High*
			Normal	*High*	*High*	*High*	*High*
	Normal		Very Low	*High*	*Middle*	*Middle*	*Middle*
			Low	*High*	*Middle*	*Middle*	*Middle*
			Normal	*High*	*Low*	*Low*	*Low*
	Late		Very Low	*High*	*High*	*Middle*	*High*
			Low	*Middle*	*Middle*	*Middle*	*Middle*
			Normal	*High*	*Middle*	*Middle*	*Low*

The risk of infant death is determined by Mamdani inference method and Centroid defuzzification method is used for crisp output value. The risks are estimated as a percentage.

Checking Validity of the Model:

The estimated model is validated by taking data from NFHS-III which contains the same variables with required information of the defined input variables. National Family Health Surveys (NFHS) of India emerged as an important source of data on population, health, and nutrition. Four rounds NFHS-I, NFHS-II, NFHS-III and NFHS-IV have been conducted till date. These surveys are conducted nationwide with a representative sample of households throughout the country. It provides estimates of fertility, family planning, infant and child mortality, reproductive and child health, nutrition of women and children, the quality of health and family welfare services and socioeconomic conditions of whole national and its states.

From the whole data base of ever married women of NFHS-III, individual cases with valid values of the input variables considered in the study together with infant mortality status are extracted. The validity of the model is tested in three stages viz.- ROC curve, appropriate statistical test(s), individual observed using extracted information and the obtained risk values of infant mortality. The MATLAB software (R2008b) and SPSS (17.0) are used to perform the analysis.

3. RESULTS AND FINDINGS

Membership Function Plots of Input and Output Variables:

The figures from 1 to 3 showed the membership function plots of inputs *viz*.-birth weight, age of the mother and anaemia (Haemoglobin in g/dl) level of the mother respectively and figure 4 showed the membership functions of output variables risk of infant mortality.

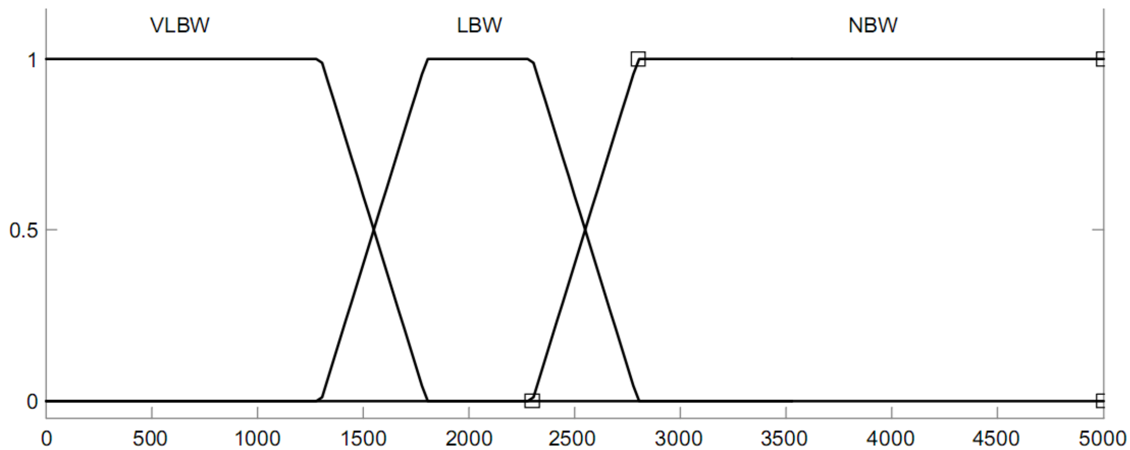

Figure1: Membership function plots of input birth weight

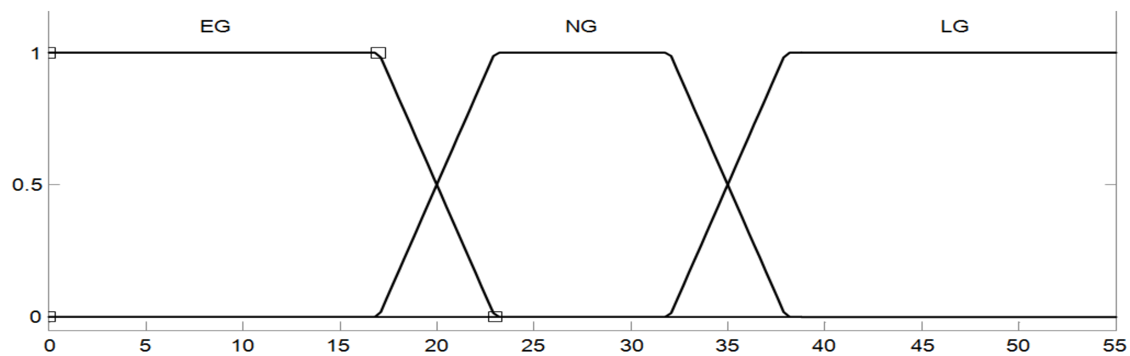

Figure 2: Membership function plots of input Mother's age

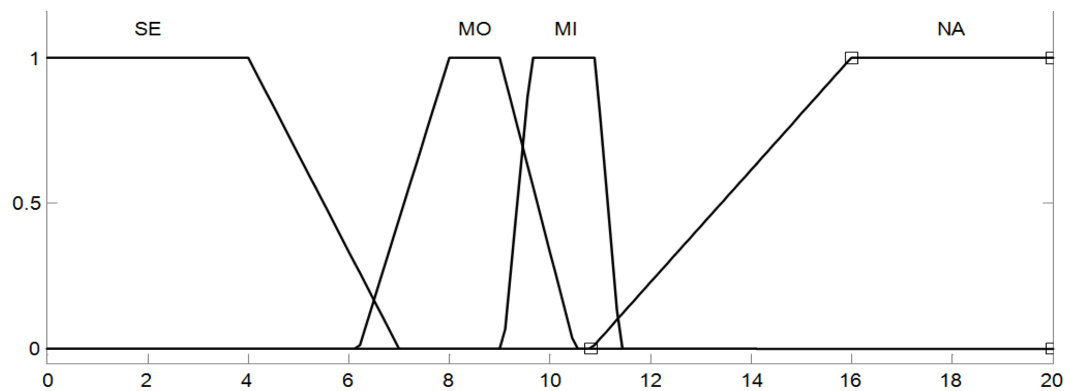

Figure 3: Membership function plots of input mother's anaemia (Haemoglobin in g/dl) level

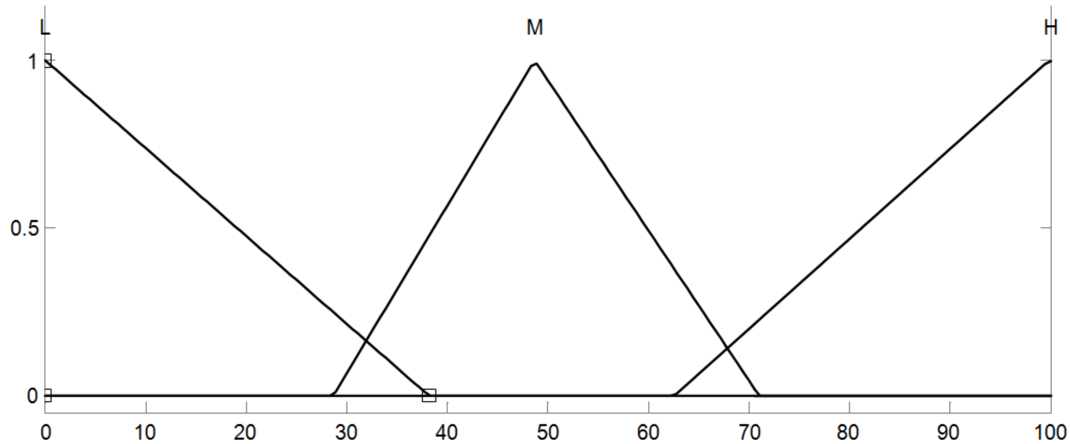

Figure 4: Membership function plots of output risk of infant mortality

Inference from the Model:

The mean of the calculated risk values is 46.35 per cent (SD = 13.13), the range of these values is 12.4-86.4 per cent and the median value was 49.6 per cent with mode 49.7 per cent.

The surface of the risk of infant death using the mother's age (in years) and birth weight (in grams); mother's anaemia level and birth weight (in grams); and mother's age (in years) and mother's anaemia are presented in figures 5, 6, 7. The risk of infant death is low when the mother's age is normal (i.e. middle age) and birth weight is normal (see Figure 5). Figure 6 shows that the risk of infant death decreases monotonically when birth weight and mother's Haemoglobin level increases. Similarly, from figure 7 it can be observed that when aneamia level and age of mother increases, the risk of infant death decreases.

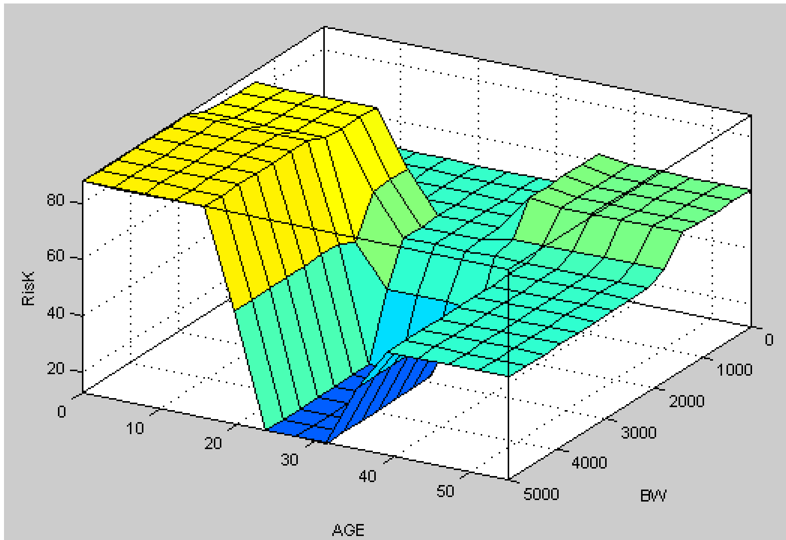

Figure 5: Surface plot of mother's age (in years) and birth weight (in grams)

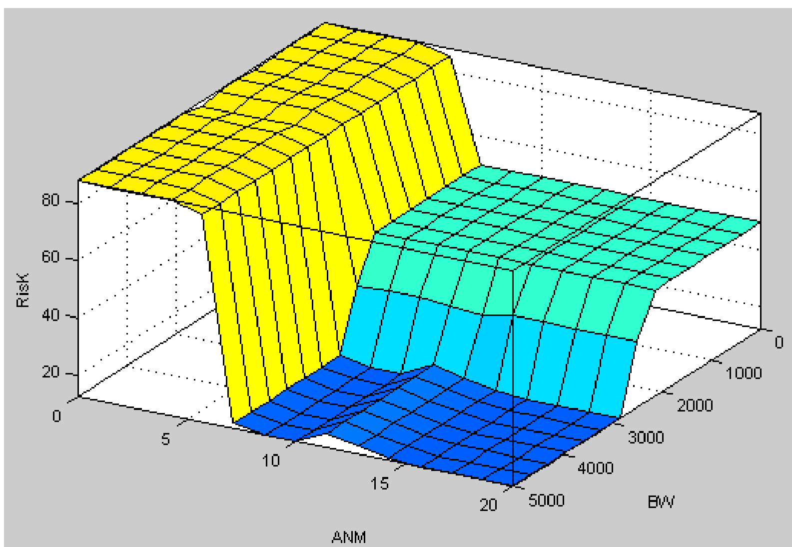

Figure 6: Surface plot of mother's anaemia and birth weight (in grams)

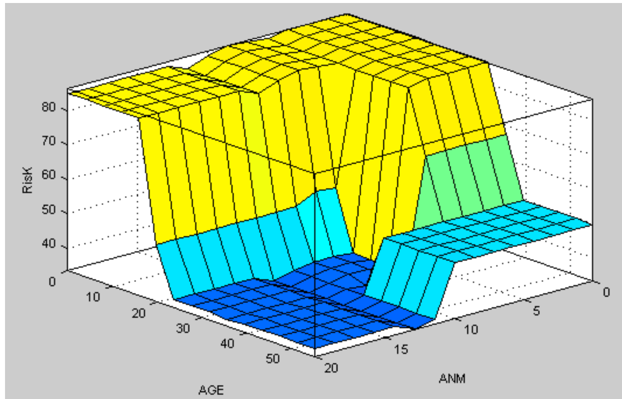

Figure 7: Surface plot of mother's aneamia and mother's age (in years)

4. VALIDITY OF THE MODEL

Receiver Operating Characteristic (ROC) Curve:

The area under the ROC curve is 0.934 (95 per cent CI: 0.84-1.00) ($p < 0.001$). Accuracy is higher when risk is below 32 per cent, corresponding to 0.91 (95 per cent CI: 0.85-0.95) in respect to sensitivity, 0.71 (95 per cent CI: 0.35-0.92) specificity, 0.99 positive predictive value and 0.26 negative predictive value. At the lowest risk 12.4 per cent, sensitivity is 1.00 (95 per cent CI: 0.97-1.00), specificity is 0.29 (95 per cent CI: 0.08-0.65), negative predictive value is 1.00, positive predictive value is 0.97 and the accuracy is 0.97. The highest accuracy (0.99) is observed at risk 17.8 per cent, corresponding sensitivity and specificity are 1.00 (95 per cent CI: 0.97-1.00) and 0.71 (95 per cent CI: 0.35-0.92) respectively with 1.00 negative predictive value, 0.99 positive predictive value.

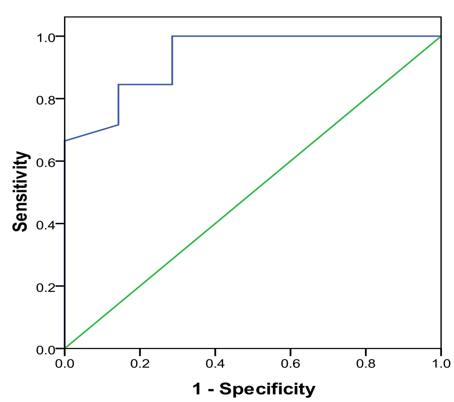

Figure 8: Receiver operating characteristic (ROC) curve

Statistical Test:

Kolmogorov-Smirnov and Shapiro-Wilk tests showed that the estimated risk values of infant mortality are not normally distributed ($p < 0.001$). The Mann-Whitney test statistic value is 3.89 with $p < 0.001$. It resulted in a mean rank of risk among infant death is 84.54; while the mean rank of risk among the survive infants is 14.29. Therefore the propose model estimate high risk value for those infants who actually died and low risk for those who survive.

Individual Observation of Cases:

By generating random numbers, 15 cases are randomly selected from NFHS-III data base which contains the information of the input variables and their outcomes i.e. dead or alive are observed individually. The last births within 5 years interval are considered for the study. The selected 15 cases with the information of input variables and output variable together with estimated risk are presented in Table 2.

Table 2: Individual observation of cases

Sl. No.	Birth Weight (in gram)	Mother's Age (in years)	Mother's Anaemia (Haemoglobin in g/dl)	Estimated Risk (in %)	Result (Dead/ Alive)
1	3,000	26	8.4	12.4	Alive
2	3,300	28	10.1	12.4	Alive
3	5,000	28	11.0	12.9	Alive
4	3,500	35	12.7	15.6	Alive
5	3,000	43	11.4	17.8	Alive
6	2,400	29	11.9	34.5	Alive
7	2,250	30	10.1	49.4	Alive
8	2,800	28	5.2	86.4	Dead
9	1,500	30	7.5	49.6	Dead
10	2,000	26	6.1	84.2	Dead
11	2,200	16	6.2	83.9	Dead
12	5,000	18	9.9	70.6	Dead
13	1,500	18	8.7	72.9	Dead
14	1,030	30	10.3	49.5	Dead
15	1,250	17	12.3	50.0	Dead

It can be observed that the fitted model assign low risk to the infants who survived than those who had died.

5. DISCUSSION

In this study fuzzy logic approach is applied to develop a fuzzy logic model to estimate the risk of infant mortality. The proposed model includes three input variables viz.- birth weight, mother's age and mother's anaemia (Haemoglobin in g/dl) level.

The developed fuzzy predictive model provides the risk of infant mortality with different levels of the input variables. For example, when birth weight of the infant is 3,300g and his/her mother's age is 28 years and his/her mother's anaemia (Haemoglobin in g/dl) level is 10.1, then the estimated risk of death for that infant is 12.4 per cent, which cannot be possible in case of logistic regression model(s).

Moreover, comparison of the fitted fuzzy logic model output i.e. risk of infant mortality with the actual data is indispensable and it determines the applicability of a model in practical situations. The ROC and statistical analysis shows that the model developed here provides good results and it pointed out its application in real fields.

6. CONCLUSION

Large number of research studies has been conducted to assess the effect of possible risk factors on infant mortality around the globe. This study used a fuzzy logic model in order to estimate the risk of infant mortality with different possible risk factors. The fitted fuzzy logic models on infant mortality are found to be valid and thus may be used in future for estimation of risk of infant mortality considering the input variables under study.

However, the fuzzy logic technique is still relatively unexplored in demography, epidemiology and public health. Typically causal analysis are the preferred outcome of statistical modeling; and logistic and linear regressions are superior to fuzzy logic in this respect as they permit for the easy recognition of associated variables with the utmost relevance in estimating results, through their related odds ratios and slop coefficients respectively. Whereas, logistic regression models use dichotomous independent variables, such as present or absent, yes or no; fuzzy logic deals with linguistic terms which are unavoidably vague and uncertain. Moreover in fuzzy logic approach each element can be classified into several categories, with different membership values. Although it has been reveal that fuzzy logic may yield superior predictive performances to conventional methods, in the area of demography, epidemiology and public health if the interest is centered on formulating fuzzy models with associated risk factor. In fuzzy logic, if research on variable selection and predictive relevance is to be widely used in these fields, these must constitute new fundamental research areas.

REFERENCES

[1] Barman, M. P. and Saikia, R. (2013). Factors influencing infant mortality in North-East India: A statistical analysis. *Scholars' view*, 1(2), 92–100.

[2] Barros, F.C. and Victora, C.G. (2008). Maternal-child health in Pelotas, Rio Grande do Sul State, Brazil: major conclusions from comparisons of the 1982, 1993, and 2004 birth cohorts. *Cad Saude Publica*, 24(3), 461–67.

[3] Benjamin, A.I., Sengupta, P. and Singh, S. (2009). Perinatal Mortality and its Risk Factors in Ludhiana: A Population-Based Prospective Cohort Study. *Health and Population: Perspectives and Issues*, 32(1), 12–20.

[4] Bora, D. and Barman, M. P. (2017). Fuzzy Linguistic Model for Estimating the Risk of Child Mortality. *International Journal of Advance Research in Computer Science and Management Studies*, 5(7), 59–67.

[5] Characteristics of Childbearing Women. Retrieved from https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&cad=rja&uact=8&ved=0ahUKEwjMqNnfj_jQAhXGuo8KHYPTAjUQFghWMAk&url=httpsper cent3Aper cent2Fper cent2Fzoek.officielebekendmakingen.nlper cent2Fblg29964.pdf&usg=AFQjCNHukSwNZN6 JXHr-23hDPvZc_-v6dg&bvm=bv.142059868,d.c2I

[6] Chaves, L. E. and Nascimento, L. F. C. (2014). Estimating outcomes in new born infants using fuzzy logic. *Rev. Paul. Pediatr*, 32(2), 164–170.

[7] Childbearing Experiences. Retrieved from https://dhsprogram.com/pubs/pdf/OD16/03 Childbearing Experiences.pdf

[8] Cooper, L. G., Leland, N. L. and Alexander, G. (1995). Effect of maternal age on birth outcomes among young adolescents. *Social Biology*, 42(1–2), 42–22.

[9] Diab, H. and Saade, J. (2005). Fuzzy Inference Modeling Methodology for the Simulation of Population Growth. *International Arab Journal of Information Technology*, 2(1), 75–86.

[10] Ezeh, O. K., Agho, K. E., Dibley, M. J., Hall, J. J. and Page, A. N. (2015). Risk factors for postneonatal, infant, child and under-5 mortality in Nigeria: a pooled cross-sectional analysis. *BMJ Open*, 5.

[11] Kalaivani, K. (2009). Prevalence and consequences of anaemia in pregnancy. *Indian J. Med. Res*, 130, 627–633.

[12] Kassar, S. B., Melo, A. M., Coutinho, S. B., Lima, M. C. and Lira, P. I. (2013). Determinants of neonatal death with emphasis on health care during pregnancy, childbirth and reproductive history. *J. Pediatr (Rio J.)*, 89(3), 269–77.

[13] Maia, L.T., Souza, W. V. and Mendes, A. C. (2012). Differences in risk factors for infant mortality in five Brazilian cities: a case-control study based on the Mortality Information System and Information System on Live Births. *Cad Saude Publication*, 28(11), 2163–76.

[14] Mathews, T. and MacDorman, M. F. (2007). Statistics NCFH: infant mortality statistics from the 2004 period linked birth, infant death data set. US Department of Health & Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

[15] Mombelli, M. A., Sass, A., Molena, C. A., Téston, E. F. and Marcon, S.S. Risk factors for infant mortality in municipalities of Paraná State. 1997-2008, Reviewed: Pediatr, P. (2012), 30(20), 187–94.

[16] Nascimento, L. F. C. and Ortega, N. R. S (2002). Fuzzy linguistic model for evaluating the risk of neonatal death. *Rev Saúde Pública*, 36(6), 686–692.

[17] Nascimento, L. F. C., Rizol, P. M. S. R. and Abiuzi, L. B. (2009). Establishing the risk of neonatal mortality using a fuzzy predictive model. *Cad. Saúde Pública Rio de Janeiro*, 25(9), 2043–2052.

[18] National Institutes of Health, (2011). Your Guide to Anemia, U.S. Department of Health and human services.

[19] Phipps, M. G., Sowers, M. and Demonner, S. M. (2002). The Risk for Infant Mortality among Adolescent Childbearing Groups. *Journal of Women's Health*, 11(10), 889–897.

[20] Rajeswari, R., Burman, B., Sundar, J. S. and Ramya, K. (2015). Trends in birth weight and the prevalence of low birth weight in a tertiary care hospital, Chennai. *Journal of Dental and Medical Sciences*, 14(8), 07–13.

[21] Sharma, J.B. and Shankar, M. (2010). Anemia in Pregnancy. *Journal of Internation Medical Sciences Academy*, 23(4), 253–260.

[22] Silva, A.L. and Mathias, T.A. (2014). Independent risk factors associated with infant deaths. *Acta Paul Enferm*, 27(1), 48–55.

[23] Smith, A. (2012). Guide to evaluation and treatment of anaemia in general practice. *Drug review Anaemia*, 25–42. Retrieved from www.prescriber.co.in

[24] Yilmaz, A. and Ayan, K. (2013). Cancer risk analysis by fuzzy logic approach and performance status of the model. *Turkish Journal of Electrical Engineering and Computer Sciences*, 21, 897–912.