Remittances Review
April 2024,
Volume: 9, No: 2, pp.4216-4226
ISSN:2059-6588(Print)|ISSN2059-6596(Online)
Received : 15 March 2024, Accepted: 09 April 2024

DOI: https://doi.org/10.33282/rr.vx9i2.219

Classifying the modes of Cesarean Section using Machine Learning Techniques

[1]Maria Malik, [2]Muhammad Awais ,
[3]Zahid Khan, [4]Ayesha Sultan, [5]Kalsoom Akhtar Chaudhry, [6]Ajab Khan

1. COMSATS university Islamabad, Pakistan Email; mariamalikkhan@gmail.com

2. Lecturer Statistics in HED Email: Ovaisqarni15@gmail.com
3. Lecturer, Department of Statistics, University of Malakand, Chakdara, Lower Dir, Khyber Pakhtunkhwa, Pakistan Email: Zahid.uom1@gmail.com

4. Lecturer, Department of Statistics, Virtual University, Islamabad, Pakistan
Email: asultan@vu.edu.pk

5. Associate Professor, Kinnaird College for Women, Lahore.
Email: kalsoom.akhtar@kinnaird.edu.pk

6. Department of Statistics, University of Malakand, Chakdara, Lower Dir, Khyber Pakhtunkhwa, Pakistan Email: Ajabstat524@gmail.com

Abstract
Machine learning techniques provide a learning method that can be used to motivate knowledge from data. There are some studies on the use of machine learning techniques for medical cesarean data. In this study, we evaluate different machine learning techniques for the general reasons for the cesarean section (emergency C-section or preplanned). Data on cesarean section is a collection from different hospitals of city Lahore and different medical factors are identified. A cesarean classification model is built using a Random Forest classifier, and K-Nearest Neighbor classifier. It can classify the cesarean into emergency C-section or preplanned C-section with an average accuracy, precision, and recall of 67% and 65% respectively. Chi_square test of association is used to extract disease patterns from the collection. It highlights the significant medical covariates that are associated with doctors' advice and the association among pre-birth covariates.

Introduction
C-section is a surgical technique, applied to save maternal and neonatal health when the normal procedure of delivery failed (Rafiei et al., 2018; Sharma & Gupta, 2018). It is one of the oldest and most commonly done surgical procedures in the field of obstetrics (Sharma & Gupta, 2018).

The Caesarean section has transformed maternity care and saved millions of women's and babies' lives. Julius Caesar was named after a forefather who was born via Caesarean section. In the early history of C-sections, it was a questionable fact that was cloaked in myth. Over time, the origins of C-sections have been proven to be incorrect. Julius Caesar's surgical birth is thought to be the source of the name (LoCicero, 2007; Hieronimus & Cortner 2016; Dregni, 2008). The phrase "Caesarean Operation" was used until the sixteenth and seventeenth centuries, when it was progressively replaced by "Caesarean Section"; the term section was introduced in 1598 by Guillimeau's book on midwifery (O'sullivan, 1990; Churchill, 1994; Low, 2009; Agarwal & Sahoo, 2019). The aim of the C-section changes from time to time as technology evolves. However, as time passed, the indications for C-sections shifted from the ancient to the modern eras (Higgins, 1985; Nobrega, 2015; Rajapreeethi, 2020). In science and medicine, the Middle Ages were mostly seen as a period of stagnation, but specific anecdotes of C-sections helped to boost and sustain trust in the operation's eventual success (Cramer, 2021). In today's world, a C-section usually takes 45 minutes to an hour to complete (Berger et al., 2011; Bernstein, 2021). After a dose of anaesthetic or a spinal block, the woman remains conscious. To drain the bladder, a urinary catheter is used, and the abdominal region is subsequently treated with antiseptic. The cut on a woman's lower abdomen is fifteen centimeters long (6 inches). After that, the uterus is opened with a second cut and the baby is delivered. It takes several days to recover from a C-section (Moore & Costa; 2003). The global rate of C-sections is gradually increasing to safeguard the mother and newborn, resulting in a significant increase in the world's population (Walsh, 2008; Morris, 2016; Rafiei, 2018). C-sections are becoming more common in both developing and developed countries around the world. Many countries have exceeded the ten to fifteen percent limit set by WHO (Wilmoth, 1995; Al Rifai, 2017; Sarkar, 2020). Pakistan is one of the developing countries where the rate of C-sections is higher than WHO recommendations (Nazir, 2015; Mumtaz, 2017; Verma et al., 2020; Nazir, 2020; Rasool et. 2021). Reasons for an operation include twin pregnancy, high blood pressure in pregnant women, wrong position of children and obstructed labor, etc. (Shabab, 2010; Sharmila, 2016; Baitieha, 2017). Rehan and Najmi (2000) elaborate on the factors in control for the high C-section rate in teaching hospitals associated with Fatima Jinnah Medical College, Lahore. They statistically signified that repeated section, dystocia, and fetal distress are risk factors for C-sections. Khawaja et al. (2004) conducted a descriptive analysis at the Obstetrics and Gynecology Unit III, Sir Ganga Ram Hospital, Lahore for a duration of six months, from September-2000 to February 2001. They analyzed the rate of C-sections was 21.07 percent. Out of 300 cesarean cases, emergency C-sections were 266 (88.67%) whereas 34 were elective (11.33%). They considered the significant indications for the emergency cesarean cases were fetal distress (22.18%), dystocia (28.20%), CPD (6.77%), and a previous C-section one (5.64%) or two C-sections (7.52%), while the significant indications in the elective C-sections were a prior two (50%) or three (2.94%) C-Sections. They also concluded the average age of the women was 29 years. We will discuss various inferences that can be drawn from the presented research. What researchers find to be the most actual set of covariates for doctors' advice is summarized in this portion? These reviews provided more knowledge on the C-section and aided in the formulation of our research's conceptual foundation. Moreover, both preexisting and gestational diabetes increase the risk for cesarean delivery, for detail of diabetes see (Akhtar et al., 2016; Khan et al. 2022a, Khan et al., 2022b; Haq et al., 2024; Haq et al., 2022). Most of the previous study used logistic regression model for assessing risk factors of the diseases, and to model

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4216-4226
ISSN:2059-6588(Print)|ISSN2059-6596(Online)

binary response variable (Khan, Hussain and Ijaz, 2022; Khan, Haq and Ali, 2022; Khan et al. 2022; Khan et al. 2022a; Khan et al. 2022b), in the present study advance techniques are sued.

Material and Methods

This is an analytical study carried out in five hospitals (Ganga Ram Hospital, Jinnah Hospital, Lady Aitcheson Hospital, Fatima Memorial Hospital, & Naz Hospital) in Lahore between January 2021 to August 2021. All the patients undergoing caesarian section. Demographic and clinical features include maternal age at the time of marriage and upon C-section, educational level, social status, category of the hospital (government/private), residential area, working status (doing the job or not), maternal weight, Reasons for C-section, maternal family history of C-section, maternal health condition, maternal short pelvic bone, baby weight at the time of caesarian, baby position, & fetal distress asked in the questionnaire. The main objective was to study the prevalence of C-sections. The Cochran's formula (1977) is used to determine the optimal sample size. A sample of 500 women was selected to gather data, the data is transformed and then analyzed by the data analysis method. The data is analyzed using SPSS.21 software and R language 4.2.0. Data is entered using the coding technique, and using SPSS.21 software to determine frequencies and percentages for descriptive variables. And also find the frequencies and percentages for categorical data. We also examine the association between variables using Pearson's Chi_square test of association. The p-value of <0.05 was taken as statistically significant.

Results and Discussion

This research is based on 500 respondents. In this case study, only females were taken into account whose baby was delivered by C-section for the year 2021. The descriptive analysis of demographic characteristics of risk factors of the C-section is given in table 1. Women who got married in the age group 15-20 years have a maximum C-section (46%). Women in the age group 21-25 & 26-30 have a maximum C-section (35%). Maternal weight also affects the cause of C-sections. Women above 70 kilograms have a maximum C-section (39%).

Table 1 Distribution (%) of maternal age upon marriage, age at C-section, and Number of Children prior to this C-section.

Age upon marriage	Age at C-Section	Weight
Year (*f*) (%)	Year (*f*) (%)	Kilograms (*f*) (%)
<15 (4) (0.80)	<15 (1) (0.20%)	<40 (2) (0.40)
15-20 (229) (45.80)	15-20 (34) (6.80)	40-50 (32) (6.40)
21-25 (180) (36.00)	21-25 (177) (35.40)	51-60 (93) (18.60)
26-30 (78) (15.60)	26-30 (177) (35.40)	61-70 (178)(35.60)
31-35 (9) (1.80)	31-35 (92) (18.40)	>71 (195) (39.00)
>35 (4) (0.80)	>35 (19) (3.80)	

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4216-4226
ISSN:2059-6588(Print)|ISSN2059-6596(Online)

Social factors (educational level, financial status, jobholder, residential area, and type of hospital) and medical factors (related to a patient health condition) proved to be the significant factors for the causes of C-sections. Table 2 depicts that most of the C-sections were conducted by women which are uneducated (25%). Women which belonged to the middle class (47%) had most of the C-section. Working women had fewer C-sections (8%). Most of the patients with C-section cases are in government hospitals (88%).

Table 2 Distribution (%) maternal educational level, Weight, Hospital (government or private), and Social status.

Educational level	**Social status**	**Doing job**	**Residential area**	**Hospital**
Level (*f*) (%)	**Status (*f*) (%)**	**Working (*f*) (%)**	**Area (*f*) (%)**	**Category (*f*) (%)**
Uneducated (127) (25.40)	Upper Class (35) (7.00)	No (460) (92.00)	Rural (115) (23.00)	Government (439) (87.80)
Middle (80) (16.00)	Middle Class (234) (46.80)	Yes (40) (8.0)	Urban (385) (77.00)	Private (61) (12.20)
Matric (118) (23.60)	Lower Class (231) (46.20)			
Intermediate (69) (13.80)				
Graduation (64) (12.80)				
Postgraduate (42) (8.40)				

Clinical factors are the actual reasons for C-sections. Table 3 depicts that out of 500 patients, 249 females had emergency C-sections and 251 females have preplanned C-sections. These factors have a natural impact on the mother's health. Most of the women have C-section cases (51%) in their families already. Women having average conditions (45%) have most of the cesarean cases.

Table 3 Distribution (%) maternal health, the reason for C-section, and anyone have C-section in her family

Type of C-section	**Does anyone in your family has a C-section**	**Health in general**
Reason (*f*) (%)	**Family member (*f*) (%)**	**Condition (*f*) (%)**
Emergency (249) (49.80)	No (244) (48.80)	Excellent (102) (20.40)
Preplanned 251) (50.20)	Yes (256) (51.20)	Good (225) (24.00)
		Average (122) (45.40)
		Poor (51) (10.20)

Moreover, Table 4 depicts an association between maternal age on marriage and age at C-section (p-value = 0.000), between no. of children with living area (0.003) educational level (p-value = 0.001), social status (0.000), maternal age and weight (0.000), social status and hospital type (0.000), family history with social status (0.028) and educational level (0.009).

Table 4 Association between Demographic Characteristics

Hypothesis	Chi_square	p-value
Association between age upon marriage & age at C-section	193.568	0.000
Association between no. of children & living area	11.436	0.003
Association between no. of children & Education level	29.785	0.001
Association between no. of children & Social status	26.827	0.000
Association between Age at C-section & Weight	68.520	0.000
Association between Social status & Hospital	63.773	0.000
Association between Social status & family history of C-section	7.151	0.028
Association between Educational level & family history of C-section	15.240	0.009

We applied the Chi-Square test to the data in order to determine whether the attributes under the research question have an association with the doctor's advice. Out of 500 cesarean patients, 280 (56%) acted upon the advice of their concerned doctor. We have chosen 5 percent as the level of significance to reject the null hypothesis.

Table 5 Association between Study Related covariates

Hypothesis	Chi_square	Sig.
Association between doctor's advice & short pelvic bone	9.671	0.002
Association between doctor's advice & baby's overweight	3.969	0.000
Association between doctor's advice and wrong position of the aby	7.316	0.000
Association between doctor's advice and fetal distress	12.384	0.000

The Chi_square tests are presented in table (5). Table (5) depicts that short pelvic bone (p-value = 0.002), baby overweight (p-value = 0.000), and wrong position of the baby (p-value = 0.000) and fetal (p-value = 0.000) are proved to be associated with doctor's advice. K-nearest neighbor classifier is used to classify the general mode of cesarean section into emergency C-section and preplanned C-section. We used the KNN package in R language 4.2.0. To evaluate the classification we have found the precision and recall values for every class i.e. emergency C-section & preplanned C-section. In the table (6) and figure 1, the precision for emergency C-sections is 58% and the recall is 69%. The precision for preplanned C-sections is 72% and the recall is 62%. This indicates that this classifier can be successfully used to predict the type of C-section. Based on these covariates. Recall, that it is likely that a relevant document will be retrieved in the search. Accuracy The probability that the recovered document is relevant. A received operator characteristic (ROC) shows how much confidence the data gives to the true positive to distinguish it from the false positive. ROC is 0.714.

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4216-4226
ISSN:2059-6588(Print)|ISSN2059-6596(Online)

Table 6 **Evaluation metrics of K-Nearest Neighbor Classifier**

Class	Precision	Recall	F-measure	ROC area
Emergency C-section	0.584	0.692	0.634	0.714
Preplanned C-section	0.722	0.619	0.667	0.714
Average	0.662	0.651	0.652	0.74

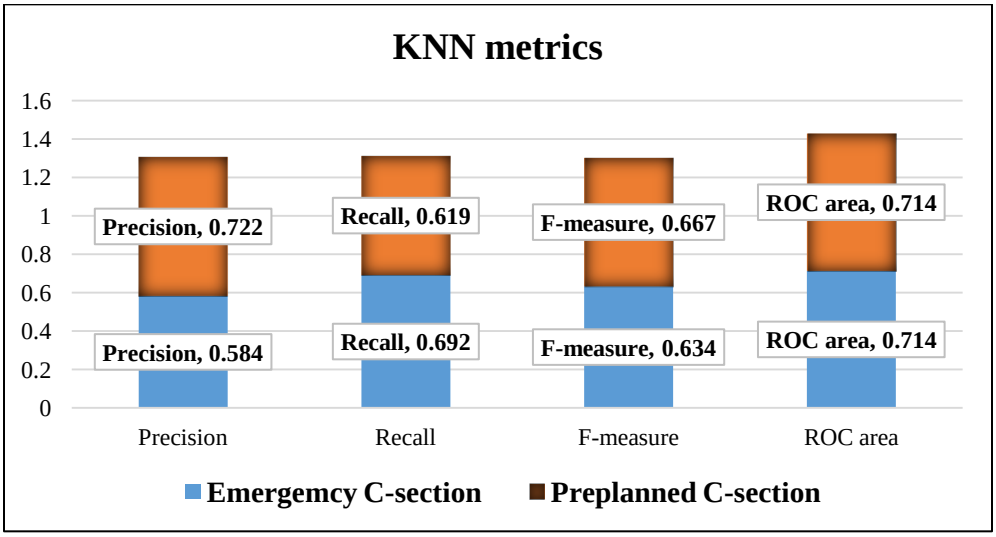

Figure 1

The second classifier is the Random Forest classifier algorithm. Then folder cross-validation is used to train and test the model. As shown in table 7 and in figure 2 the precision and recall for emergency C-sections and preplanned-section is 63% and 74% respectively. For preplanned cesarean class, the precision and recall are 72% and 62% respectively. The precision and recall values indicate that the Random Forest classifier can be successfully used to predict the type of cesarean. The F-measure is 0.667 in the given data. ROC area is 0.76 in calculated data.

Table 7 **Evaluation metrics of Random Forest Classifier**

Class	**Precision**	**Recall**	**F-measure**	**ROC area**
Emergency C-section	0.639	0.746	0.688	0.757
Preplanned C-section	0.727	0.615	0.667	0.77
Average	0.685	0.678	0.667	0.763

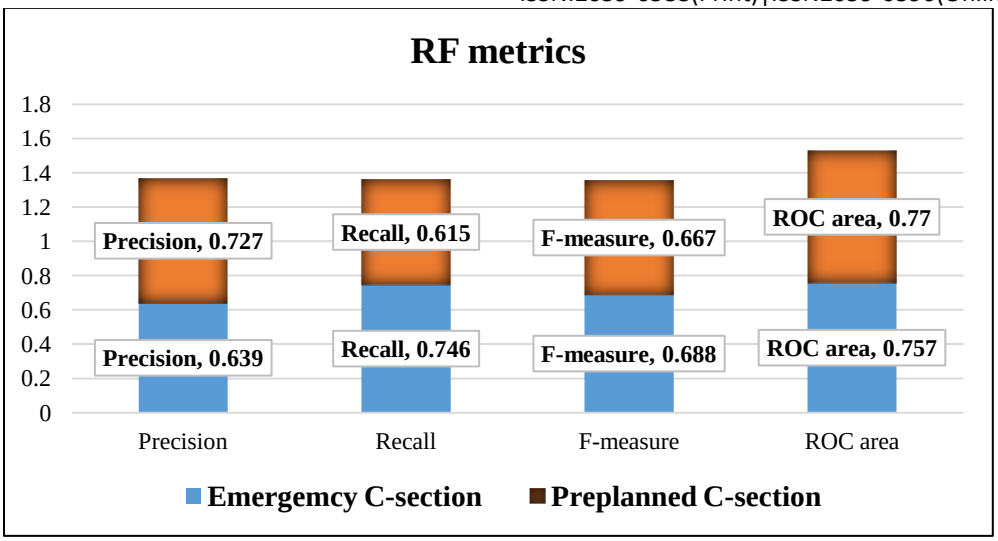

Figure 2

Conclusion

In this article, we present the successful use of machine learning techniques in the medical domain. Knowledge engineering and machine learning are used to extract disease samples from available medical data. The extracted patterns can also be used for medical diagnosis, prognosis, and treatment. This paper has covered three goals: First, it has identified the significant factors that influence the reasons for C-sections. Second, it has presented a prediction model for the reason of birth by C-section that can help both the doctors and patients. Third, it has presented the association between doctor's advice & medical factors and the association among pregnancy covariates. K-nearest neighbor classifier can classify between emergency cesarean and preplanned cesarean with 65% accuracy. Random Forest classifier can classify between emergency cesarean and preplanned cesarean with 67% accuracy. The Chi-square test indicates a significant association between doctor's advice and clinical covariates of maternal including short pelvic bone, baby overweight, baby wrong position, and fetal distress. Similarly, maternal demographic characteristics are associated with and affected by the reasons for cesarean section. Briefly, this paper has highlighted the usefulness of machine learning classification techniques in the medical domain.

Future Work

We have obtained quite satisfactory results for the classification of general reasons for the cesarean section. However, these results can be further improved by the identification of additional covariates that influence whether the cesarean either be preplanned or emergency. We will investigate these covariates in the future. As two machine learning techniques have been used in this research work. We can evaluate more machine learning classification techniques on the cesarean data. The prediction models will be more useful if we apply more techniques. We want to increase the area surveyed because there may be certain geographical factors that can influence the type of birth.

Agrawal, R., & Sahoo, G. (2019). Evolution of Cesarean Section. *Journal of Medical Science And clinical Research*, *7*(04), 193-201.

Al Rifai, R. H. (2017). Trend of caesarean deliveries in Egypt and its associated factors: evidence from national surveys, 2005–2014. *BMC pregnancy and childbirth*, *17*(1), 1-14.

Batieha, A. M., Al-Daradkah, S. A., Khader, Y. S., Basha, A., Sabet, F., Athamneh, T. Z., & Sheyyab, M. (2017). Cesarean section: incidence, causes, associated factors and outcomes: a national prospective study from Jordan. *Gynecol Obstet Case Rep*, *3*(3), 55.

Berger, H., Sistonen, J., Madadi, P., Matok, I., Gijsen, V. M., de Wildt, S. N., ... & Koren, G. (2011). CYP2D6 polymorphisms and codeine analgesia in postpartum pain management: a pilot study. *Therapeutic drug monitoring*, *33*(4), 425-432.

Bernstein, J., Spitzer, Y., Ohaegbulam, K., Reddy, S., Song, J., Romanelli, E., & Nair, S. (2022). The analgesic efficacy of IV acetaminophen for acute postoperative pain in C-section patients: a randomized, double-blind, placebo-controlled study. *The Journal of Maternal-Fetal & Neonatal Medicine*, *35*(5), 933-940.

Churchill, H. (1994). *Caesarian birth: conflict in maternity services* (Doctoral dissertation, Middlesex University).

Cramer, R. A. (2021). *Birthing a movement: Midwives, law, and the politics of reproductive care*. Stanford University Press.

Dregni, M. (2008). *Gypsy Jazz: In Search of Django Reinhardt and the Soul of Gypsy Swing*. Oxford University Press

Rafiei, M., Ghare, M. S., Akbari, M., Kiani, F., Sayehmiri, F., Sayehmiri, K., & Vafaee, R. (2018). Prevalence, causes, and complications of cesarean delivery in Iran: A systematic review and meta-analysis. *International journal of reproductive biomedicine*, *16*(4), 221.

Sharma, K., & Gupta, M. (2017). Caesarean Section: A technical update. *AOGD BULLETIN*, *17*(4), 9-12.

Sharmila, G., & Nishitha, C. (2016). Study of primary caesarean section in multigravida. *Asian pacific journal of health sciences*, *3*(4), 89-94.

Shabab, U., & Tahir, S. (2010). Effect of obesity on cesarean section rate. *Journal of Surgery Pakistan (International)*, *15*(2), 92.

Low, J. (2009). Caesarean section—past and present. *Journal of obstetrics and gynaecology Canada*, *31*(12), 1131-1136.

O'sullivan, J. F. (1990). Caesarean birth. *The Ulster medical journal*, *59*(1), 1.

Higgins, C. S. (1985). *CESAREAN SECTION: DO ECONOMIC INCENTIVES MATTER?(AN ANALYSIS OF THE IMPACT OF ECONOMIC INCENTIVES ON THE RATE OF C-SECTION AT THE HOSPITAL LEVEL)(DELIVERIES, OBSTETRICS, BIRTHING, MALPRACTICE, HMO)* (Doctoral dissertation, University of Maryland, Baltimore County).

Nobrega, J. K. (2015). *What is Pushing Brazil Not to Push?* (Doctoral dissertation, University of Illinois at Chicago).

Rajapreeethi, M. (2020). *A Prospective Observational study of Maternal and Perinatal Complications in Subsequent Pregnancy following Previous Caesarean in Tertiary Care Center* (Doctoral dissertation, Sree Mookambika Institute of Medical Sciences, Kulasekharam).

.

(Reducing Caesarean Section Rates- A multifaceted challenge).

LoCicero, D. (2007). *Superheroes and gods: A comparative study from Babylonia to Batman*. McFarland.

Hieronimus, R., & Cortner, L. E. (2016). *The Secret Life of Lady Liberty: Goddess in the New World*. Simon and Schuster.

Moore, M. C., & de Costa, C. M. (2003). *Cesarean section: Understanding and celebrating your baby's birth*. JHU Press.

Morris, T. (2016). *Cut it out: The C-section epidemic in America*. NYU Press.

Rafiei, M., Ghare, M. S., Akbari, M., Kiani, F., Sayehmiri, F., Sayehmiri, K., & Vafaee, R. (2018). Prevalence, causes, and complications of cesarean delivery in Iran: A systematic review and meta-analysis. *International journal of reproductive biomedicine*, *16*(4), 221.

Walsh, J. A. (2008). Evolution & the cesarean section rate. *The American Biology Teacher*, *70*(7), 401-404.

Sarkar, S. (2020). Prevalence and determinants of the use of caesarean section (CS) in the dichotomy of 'public'and 'private'health facilities in West Bengal. India. *Clinical Epidemiology and Global Health*, *8*(4), 1377-1383.

Wilmoth, T. A., & Elder, J. P. (1995). An assessment of research on breastfeeding promotion strategies in developing countries. *Social science & medicine*, *41*(4), 579-594.

Mumtaz, S., Bahk, J., & Khang, Y. H. (2017). Rising trends and inequalities in cesarean section rates in Pakistan: Evidence from Pakistan Demographic and Health Surveys, 1990-2013. *PloS one*, *12*(10), e0186563.

Verma, V., Vishwakarma, R. K., Nath, D. C., Khan, H. T., Prakash, R., & Abid, O. (2020). Prevalence and determinants of caesarean section in South and South-East Asian women. *Plos one*, *15*(3), e0229906.

Nazir, S., & Cready, C. (2020). The C-Section Epidemic in Pakistan. *PIDE Blog*.

Nazir, S. (2015). *Determinants of cesarean deliveries in Pakistan*. Islamabad, Pakistan: PIDE.

Rasool, M. F., Akhtar, S., Hussain, I., Majeed, A., Imran, I., Saeed, H., ... & Alqhtani, H. (2021). A cross-sectional study to assess the frequency and risk factors associated with cesarean section in Southern Punjab, Pakistan. *International Journal of Environmental Research and Public Health*, *18*(16), 8812.

Rehan, R. S. N. N. (2000). Prevalence and determinants of caesarean section in a teaching hospital of Pakistan. *Journal of Obstetrics and Gynaecology*, *20*(5), 479-483.

Khawaja, N. P., Yousaf, T., & Tayyeb, R. (2004). Analysis of caesarean delivery at a tertiary care hospital in Pakistan. *Journal of obstetrics and gynaecology*, *24*(2), 139-141.

Khan, Z., & Ali, A. (2022). Anemia and its Causes among Children in District Dir Lower Applying Ranked Set Sampling Method for Data Collection. *Research Journal of Social Sciences and Economics Review*, *3*(4), 34-37.

Khan, Z., Haq, N., & Ali, A. (2022). Job Satisfaction among University Faculty in Khyber Pakhtunkhwa Pakistan. *sjesr*, *5*(4), 1-4.

Khan, Z., Hussain, J., Ali, A., & Ali, S. (2022b). Factors Affecting Students Perception towards Teaching Faculty Evaluation at University of Malakand. *Journal of Policy Research*, *8*(2), 115-118.

Khan, Z., Hussain, S. & Ijaz, M. (2022). Assessing the effect of campus support facilities sports programs experiences provided and environmental factors on the student's satisfaction level using ordinal logistic regression model, *Journal of Positive School Psychology*, *6*(9), 71-75.

Khan, Z., Ismail, M., Hussain, J., & Azeem, M. (2022). The Impact of Salary Job Security and Workload on Teachers Satisfaction in Private Schools. *Journal of Policy Research*, *8*(3), 278-280.

Khan, Z., Israr, M., Hashim, S. and Khattak, S. W. (2022a). Perception Of the Community Toward Female Education in Lower Dir Khyber Pakhtunkhwa, *Indian Journal of Economics and Business,* 21(2), 361-364.

Khan, Z., ul Haq, B., & Amin, A. (2022). Prevalence and Modelling of Complication of Diabetes and Its Risk Factors in District Dir Lower Khyber Pakhtunkhwa. *Journal of Educational Research and Social Sciences Review (JERSSR)*, *2*(3), 88-92.

Akhtar, S., Khan, Z., Rafiq, M., & Khan, A. (2016). Prevalence of type II diabetes in District Dir Lower in Pakistan. *Pakistan journal of medical sciences*, *32*(3), 622.

Haq, B. U, Khan, Z., Khan, S., Ullah, U., Rashid, H. U. (2024). Prevalence And Associated Risk Factors of Stress Among the Undergraduate University Students in Khyber Pakhtunkhwa Province Pakistan, *Journal of Population Therapeutics & Clinical Pharmacology*, 31(1), 2771-2782

Haq, B. U., Ullah, U., & Khan, Z. (2022). Prevalence of Stress and its Risk Factors among the Students in University of Malakand, Khyber Pakhtunkhwa, Pakistan. *Journal of Educational Research and Social Sciences Review (JERSSR)*, *2*(4), 32-38.