BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Predictive model for congenital heart disease in children of Pakistan by using structural equation modeling

Sana Shahid[1], Haris Khurram[2,3], Muhammad Ahmed Shehzad[1*] and Muhammad Aslam[4*]

Abstract

Background The structural abnormality of the heart and its blood vessels at the time of birth is known as congenital heart disease. Every year in Pakistan, sixty thousand children are born with CHD, and 44 in 1000 die before they are a month old. Various studies used different techniques to estimate the risk factors of congenital heart disease, but these techniques suffer from a deficiency of capacity to present human understanding and a deficiency of adequate data. The current study provided an innovative approach by defining the latent variables to handle this issue and building a reasonable model.

Method Data used in this study has been collected from mothers and hospital records of the children. The dataset contains information on 3900 children who visited the OPD of the Chaudry Pervaiz Elahi Institute of Cardiology (CPEIC) Multan, Pakistan from October 2021 to September 2022. The latent variables were defined from the data and structural equation modeling was used to model them.

Result The results show that there are 53.6% of males have acyanotic CHD and 54.5% have cyanotic CHD. There are 46.4% of females have acyanotic CHD and 45.5% have cyanotic CHD. The children who have no diabetes in the family are 64.0% and children who have diabetes in the family are 36.0% in acyanotic CHD, the children who have no diabetes in the family are 59.7% and children have diabetes in the family are 40.3% in cyanotic CHD. The value of standardized root mean residual is 0.087 is less than 0.089 which shows that the model is a good fit. The value of root mean square error of approximation is 0.113 is less than 0.20 which also shows the good fit of the model.

Conclusion It was concluded that the model is a good fit. Also, the latent variables, socioeconomic factors, and environmental factors of mothers during pregnancy have a significant effect in causing cyanotic while poor general health factor increases the risk of Acyanotic congenital heart disease.

Keywords Congenital heart disease, Root mean square error of approximation, Standardized root mean square residual, Structural equation modeling

*Correspondence:
Muhammad Ahmed Shehzad
ahmad.shehzad@bzu.edu.pk
Muhammad Aslam
aslam_ravian@hotmail.com
[1]Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan
[2]Department of Mathematics and Computer Science, Faculty of Science and Technology Prince of Songkla University, Pattani Campus, Thailand
[3]Department of Sciences and Humanities, National University of Computer and Emerging Sciences, Chiniot-Faisalabad Campus, Faisalabad, Pakistan
[4]Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia

Introduction

In children, congenital heart disease (CHD) is the main cause of morbidity and mortality all over the world [1, 2]. The structural abnormality of the heart and its blood vessels at the time of birth is known as CHD [3]. The CHD incidence was found at 17.9 per 1000 live births all around the world, while it was found at 16.6 per 1000 in females and 19.1 per 1000 in males [4]. Every year in Pakistan, sixty thousand children are born with CHD, and 44 in 1000 die before they are a month old [5]. Among Pakistani infants, the prevalence of CHD was found to be 4 per 1000 live births [6]. In newborns, the prevalence of congenital heart defects (CHD) was found to be higher compared to children in older age groups. In rural areas of Pakistan, the prevalence was observed to be 3.4 per 1000 live births [7]. In both children and adult populations, CHD is a significant cause of mortality, with a prevalence rate of 4 per 1000 live births in Karachi, Pakistan [8]. CHD development is associated with the mother's health condition, genetic disorders, and environmental factors [9, 10]. Identifying the factors that cause CHD in children can help prevent deaths by enabling doctors and medical practitioners to provide appropriate treatment at the right time. Many research studies aim to reduce mortality and morbidity in children. Despite variations in existing approaches, some commonalities are found to be possible, such as the selection of variables as risk factors for CHD in children. In various studies, variables such as smoking [11], family history [12], family history of diabetes, poor nutrition [13], low income, parents' education [14], dwelling area, home environment [15, 16], health care access [17, 18], and the quality of health care facilities and housing conditions [19, 20] have been identified.

Various studies have employed different techniques to estimate the risk factors of CHD. However, these techniques suffer from a lack of capacity to convey a clear understanding of the issue and a deficiency of adequate data. To the best of our knowledge, no one has utilized the Structural Equation Model (SEM) to construct a predictive model for CHD or to identify latent risk factors.

In the current study, our objective is to develop a predictive model for CHD using categorical SEM, which will also identify the latent risk factors of CHD in children.

Methodology

The current cross-sectional study is based on CHD data collected from Chaudry Pervaiz Elahi Institute of Cardiology (CPEIC) in Multan, Pakistan, from October 2021 to September 2022. We considered a sample of 3900 children who visited the hospital's outpatient department (OPD) and were diagnosed with CHD during the mentioned period. Data related to all factors were obtained from the mothers of the CHD patients and hospital records. We collected data from all eligible children present at the hospital during the specified period. Only children with confirmed CHD diagnoses via echocardiography were included in the study, while those with comorbidities were excluded from consideration.

The inclusion and exclusion criteria for the children is follow, the children have conformed diagnosis of CHD as verified by echocardiography and if any comorbid medical conditions are presence in children, they were excluded.

Data collected from mothers pertained to factors related to their pregnancy period. We developed a Performa to collect data from respondents.

The sample size determination is done by using the following formula [21].

$$n = \frac{z^2 p\,(1-p)}{\varepsilon^2} \times D$$

,where z is the table value standard normal distribution for a defined significance level and p denotes the disease prevalence because we are interested in the presence of the disease. ε denotes the level of precision and fixed within 10% and D is the effect of sampling design whose value lies between 1 and 10. For a simple random sampling, D is chosen to be 1 while for purposive sampling, the value of D is equal to 10 [22]. The CHD prevalence used for the sample size is 4 per 1000 live births.

In this research, the questionnaire was designed (Supplementary file S1) by the principal author in consultation with an expert and senior Pediatric cardiologist. This self-designed questionnaire included questions related to socioeconomic factors, environmental factors, and general health for data collection. Family history of diabetes was inquired from the parents' first-degree relatives. Family history of heart disease was inquired from the parents' first-degree relatives. Mothers who could walk at least two and a half hours a week during pregnancy were considered physically active [23, 24]. The education of parents was categorized into five groups based on formal education: uneducated, primary to middle, secondary to higher secondary, graduate, and master's or higher. The father's occupation was divided into five groups: those who were not working or deceased (Group 1), laborers or former (Group 2), those with private jobs (Group 3), small business owners (Group 4), and civil servants (Group 5). The number of children living in a home. Access to healthcare facilities was defined by the presence of a government hospital or medical unit and government doctor in their area. Good quality of health care facilities was defined by the presence of well-trained and motivated staff, accurate medical records, access to water, proper hand hygiene, energy supply, functional waste disposal facilities, reliability, safety, sanitation, and adequate stocks of medicines, supplies, and equipment.

The dependent variable was categorized into cyanotic and acyanotic CHD. It was written in the patient's medical record that the disease was cyanotic or acyanotic CHD. Doctors followed a systematic approach to diagnose after a thorough examination and documented their findings in the patient's file.

To assess the validity of the questionnaire, a pre-test was conducted. This method proved valuable in refining the questionnaire by removing certain factors and adding others to enhance data collection. During the pre-test, data was gathered from a sample of 80 children. After data collection, we performed descriptive statistics and ultimately removed questions that were deemed unnecessary for children, such as waist-hip ratio and blood group.

All the data were collected using a questionnaire that included the aforementioned variables, and for the statistical analysis, these variables were pre-coded. Both descriptive and analytical analyses were conducted on the data using the R programming language.

The primary concern in data collection through interviews is addressing bias efficiently. In this study, uniform questions were posed to each respondent to minimize the potential for bias. Several steps were taken to reduce bias in data collection, including:

1. Before conducting interviews, all targeted children and their parents received a comprehensive briefing about the study's purpose.
2. Questions were presented in simple and understandable language, and any doubts regarding the study were clarified.
3. Adequate time was allocated to each subject for the interview, with a guarantee of record confidentiality.
4. The first author personally conducted interviews with all targeted children and their parents.

Mothers were asked to provide information about socioeconomic and environmental factors based on their past experiences. Furthermore, they discussed the situation with their husbands and other family members (due to the joint family system) to minimize recall bias. However, future follow-up studies may be conducted to further refine recall bias reduction.

Table 1 Latent factors and their observed variables

No.	Latent Factor	Description	Observed variables
1	SF	Socioeconomic Factors	Education level of mother, Education level of father, Occupation of father
2	EF	Environmental Factors	Healthcare quality, access to healthcare facilities
3	GH	General Health	Diabetes in family, Family History, Physically active mother

Structural equation modeling

The combination of multiple regression and factor analysis is referred to as the Structural Equation Model (SEM) [25], which involves both measured variables and the development of latent factors from those variables. The primary objective is to model the dependent variable using latent variables. In this study, we utilized categorical SEM to model CHD in a single comprehensive model and identify hidden or latent variables that represent potential risk factors for CHD. The response variable was categorical, and the 10 observed variables were grouped into three latent factors.

For the SEM model, the 3 categories are used as latent variables defined in Table 1.

Thus, the final latent variable regression model is as follows:

$$Disease = \beta_0 + \beta_1 SF + \beta_2 EF + \beta_3 GF + \in$$

Where the disease is a categorical variable representing the Cyanotic and Acyanotic for CHD. β_i are the parameters of the latent variables and $\in$ is the independently and identically distributed error term.

Performance evaluation

To evaluate the performance of SEM, we utilize measures such as Chi-Square, Akaike information criteria, Bayesian information criteria, Standardized Root Mean Square Residual (SRMR), and Root Mean Square Error of Approximation (RMSEA). More details can be seen in [26].

The introduction of SRMR is necessary because the root mean square residual depends on the size of variances and covariances. Without considering the scales of variables, it's not possible to determine whether a given RMSEA value indicates a good or bad fit. A zero value of SRMR signifies a perfect fit. However, due to its sensitivity to misspecified models and its dependence on sample size, it is challenging to establish a definitive cutoff point for what constitutes an acceptable or good fit [27]. A rule of thumb for a good fit is that SRMR should be less than 0.05 [28]. However, the value of SRMR smaller than 0.10 will be acceptable [29].

The RMSEA was bounded below zero. It was defined as a close fit as its value was less than or equal to 0.05 [30, 31]. The value of RMSEA greater than 0.10 is not acceptable, if this value lies between 0.08 and 0.10 it can be considered a mediocre fit, if this value lies between 0.05 and 0.08 it can be considered an adequate fit, and it can be considered a good fit if the value of RMSEA less than or equal to 0.05 [31]. The RMSEA value for a good model should be less than 0.05 and the value of RMSEA less than 0.06 is considered a cutoff criterion [32]. Also,

it was observed as relatively independent of sample size [31, 33].

Results

The results of univariate analyses, including frequencies and percentages, are presented in Table 2. According to the table, 53.6% of males have acyanotic CHD, while 54.5% have cyanotic CHD. For females, 46.4% have acyanotic CHD, and 45.5% have cyanotic CHD. In the case of children with no family history of diabetes, 64.0% have acyanotic CHD, and 36.0% have cyanotic CHD. Meanwhile, for children with a family history of diabetes, 59.7% have acyanotic CHD, and 40.3% have cyanotic CHD. Lastly, children with no family history are at 31.7% for acyanotic CHD and 68.3% for cyanotic CHD, while children with a family history are at 42.9% for acyanotic CHD and 57.1% for cyanotic CHD.

Of the mothers of children having acyanotic CHD, 55.1% are uneducated, 37.0% are primary or middle educated, secondary 7.4% are higher secondary educated, 0.3% are graduates and 0.3% are masters or higher educated. Of the mother of children having cyanotic CHD 58.0% are uneducated, 33.9% are primary or middle educated, secondary and 7.6% are higher secondary educated, and 0.5% are graduated. Of the father of children having acyanotic CHD 40.7% are uneducated, 45.0% are primary or middle educated, secondary and 12.0% are higher secondary educated, 1.9% are graduate and 0.4% are masters or higher educated. Of the mother of the children having cyanotic CHD 59.0% are uneducated, 29.7% are primary or middle educated, secondary and 9.9% are higher secondary educated, 1.0% are graduated and 0.3% are masters or higher educated. Of the father of children having acyanotic CHD, 0.1% are dead or unemployed, 68.9% are labor or former, 7.2% have private jobs, 22.9% have a small business, and 1.0% have a civil servant. Of the fathers of children having cyanotic CHD, 0.3% are dead or unemployed, 69.4% are labor or former, 1.7% have a private job, 27.6% have a small business, and 1.0% have a civil servant. In acyanotic CHD, 56.8% of children have a poor quality of basic healthcare facilities, 24.4% of children have a normal quality of basic healthcare facilities, and 18.8% of children have a good quality of basic healthcare facilities. In cyanotic CHD, 42.5% of children have a poor quality of basic healthcare facilities, 34.8% of children have normal quality of basic healthcare facilities, and 22.7% of children have good quality basic healthcare facilities. In acyanotic CHD, 75.1% of children have no access to basic healthcare facilities while 24.9% of children have access to basic healthcare facilities. In cyanotic CHD, 64.7% of children have no access to basic healthcare facilities while 35.3% of children have access to basic healthcare facilities.

The more powerful alternative to multiple regression analysis is SEM. The SEM has represented a path model that permits to estimate of the effect that is direct or indirect.

In Fig. 1, an ellipse represents the latent variables, the rectangles represent the observed variables. The parameters are represented by arrows to be estimated with values and names displayed above each edge.

Table 3 shows that all the observed variables are significantly contributing to forming the latent variable at a 1% level of significance. And hence are the risk factors. Similarly, the results of the latent factor show that socioeconomic factors (SE) and environmental factors (EF)are statistically significant at a 1% level of significance. Hence these latent factors are highly associated with risk factors of CHD in children.

In Table 4 results of model performance measures are present, and the p-value of the chi-square statistic is found significant which shows that the model is a

Table 2 Descriptive analysis of the variables of CHD

Variable	Category	Acyanotic CHD	Cyanotic CHD	Variable	Category	Acyanotic CHD	Cyanotic CHD
Gander	Female	1258 (46.4%)	518 (45.5%)	Inactive	No	800 (29.5%)	342 (28.7%)
	Male	1452 (53.6%)	672 (54.5%)		Yes	1910 (70.5%)	848 (71.3%)
Diabetes	No	1734(64.0%)	710 (59.7%)	Family History	No	858 (31.7%)	510 (42.9%)
	Yes	976 (36.0%)	480 (40.3%)		Yes	1852 (68.3%)	680 (57.1%)
Mother Education	Uneducated	1492 (55.1%)	690 (58.0%)	Father Education	Uneducated	1102 (40.7)	702 (59.0)
	Primary/Middle	1002 (37.0%)	404 (33.9%)		Primary/Middle	1220 (45.0)	354 (29.7)
	Secondary/Higher	200 (7.4%)	90 (7.6%)		Secondary/Higher	326 (12.0)	118 (9.9)
	Graduate	8 (0.3%)	6 (0.5%)		Graduate	52 (1.9)	12 (1.0)
	Masters or higher	8 (0.3%)	0 (0.0%)		Masters or higher	10 (0.4)	4 (0.3)
Health care Access	No	2034 (75.1%)	770 (64.7%)	Father Occupation	Dead/ Unemployed	4 (0.1%)	4 (0.3%)
	Yes	676 (24.9%)	420 (35.3%)		Labour/Former	1866 (68.9%)	826 (69.4%)
Health Care Quality	Poor	1540 (56.8%)	506 (42.5%)		Private Job	194 (7.2%)	20 (1.7%)
	Normal	660 (24.4%)	414 (34.8%)		Small Business	620 (22.9%)	328 (27.6%)
	Good	510 (18.8%)	270 (22.7%)		Civil Servant	26 (1.0%)	12 (1.0%)

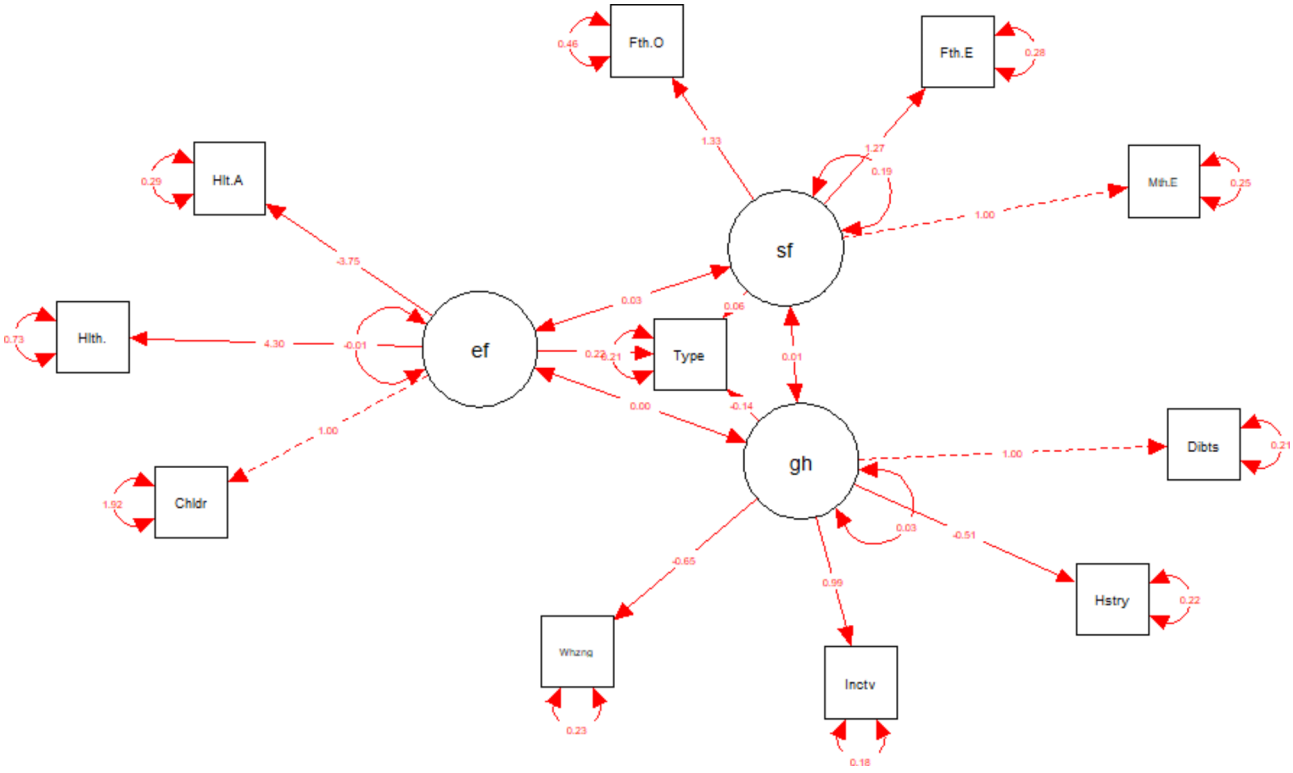

Fig. 1 The SEM model for CHD

Table 3 Regression result using SEM for CHD

Latent Variables	Observed Variables	Estimate	S.E	Z-value	*P*-value
SF		0.062	0.021	2.992	0.003**
EF		0.221	0.063	3.503	0.000***
GH		-0.144	0.088	-1.633	0.102
SF	Mother Education	1			
	Father Occupation	1.326	0.042	31.721	0.000***
	Father Education	1.270	0.037	34.049	0.000***
EF	Health Care Quality	4.302	0.693	6.205	0.000***
	Health Care Access	-3.753	0.618	-6.075	0.000***
GH	Diabetes	1			
	History	-0.505	0.100	-5.030	0.000***
	Inactive	0.986	0.152	6.490	0.000***

*** = significance at 1%; ** = significant at 5%;. S.E = Standard Error

Table 4 The model performance measures of SEM

Model Fit Test Statistic	1973.911
Degrees of freedom	39
P-value (Chi-square)	0.000
Akaike (AIC)	76132.199
Bayesian (BIC)	76215.661
RMSEA	0.113
SRMR	0.087

good fit. The values of AIC and BIC are 76132.199 and 76215.661 respectively. The value of SRMR is 0.087 is less than 0.089 which shows that the model is a good fit. The RMSEA value is 0.113, less than 0.20 which also shows the good fit of the model.

Discussion

With the help of SEM, the current study is showing a method for categorical data analysis to recognize the robustness of the relationship between associated variables to the cause of CHD in children. The importance of the SEM is that this method does not depend directly on the judgment of skilled users to build it. With the help of SEM, the best assessment model results are obtained. By using the distinct data set we could need several adjustments by applying the given approach here, i.e., a model calibration requires to be executed once again. For further studies, the analysis of the categorical data through SEM remains a challenge.

In socio-economic factors, the education and occupation of the father are found significant variables for CHD. In a previous study [14]education of parents was also found significant. Environmental factors, poor access to healthcare facilities, and poor quality of healthcare facilities are found to be significant factors for CHD. The same variables were significant in studies [17–19]. In general health factors, the family history of heart and physically inactive mothers found significant variables for CHD,

and the family history was also found significant in the study [12]. Studies [12, 14, 17–19] support our findings related to the observable factor, but the latent factor was not explored in any of the studies for CHD.

In the current study, 1452 males had acyanotic CHD and 672 had cyanotic CHD. There are 1258 of females have acyanotic CHD and 518 have cyanotic CHD. Similar results were found in another study conducted in India [32] and their results show that 77 of males have Acyanotic CHD and 33 have cyanotic CHD. There were 64 females have Acyanotic CHD and 27 have cyanotic CHD. In two related studies done in Africa [34] in Cameroon and in Nigeria [35], the proportion of acyanotic heart disease was well above 80%.

The result of the SEM from another study [36] for coronary heart disease shows that the education of parents, the status of employment, and family history were the contributing factors to coronary heart disease. The value of chi-square was 0.966 while the value of RMSEA was 0.000 found in this study. Another study [37, 38] shows that the low level of education of mothers was a significant risk factor for coronary heart disease. A low level of education may cause insufficient knowledge about birth and maternal outcomes including the health of the mother and physical activity. Another study [39] reports that the risk of CHD was linked to genetic, behavioral, and environmental factors, Also the development of CHD is associated with a low level of education of the mother, the occupation of the parent, and the status of employment.

Conclusion

In this study, categorical SEM is employed to develop a predictive model for CHD and identify latent risk factors in children. Diagnostic measures assessing the model's performance indicate that it is a good fit. Additionally, the latent variables, specifically socioeconomic and environmental factors during mothers' pregnancies, exhibit a significant impact in causing cyanotic congenital heart disease, while a poor general health factor increases the risk of acyanotic congenital heart disease. Among the associated observable factors, low paternal education, paternal occupation, lack of access to and poor quality of basic healthcare facilities, family history, and maternal physical inactivity during pregnancy emerge as significant risk factors. These results provide valuable insights for medical practitioners and scientists, enhancing their understanding of disease behavior and incidence, and enabling prediction and prevention strategies for CHD in children. Moreover, the study sheds light on hidden latent risk factors for the disease, which can inform effective policies and raise awareness of the causes of CHD in children. Ultimately, this knowledge contributes to the reduction of the burden of CHD-related mortality and morbidity in children.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-024-02774-y.

Supplementary Material 1

Acknowledgements
No.

Author contributions
S.S., and H.K conceptualize the study and collect the data. S.S., H.K., M.A.S, and M.A analysis, wrote and revised the final draft.

Funding
None.

Data availability
The data can be available from Sana Shahid (sanashahid052@gmail.com) and Haris Khurram (hariskhurram2@gmail.com) upon reasonable request.

Declarations

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Ethics approval and consent to participants
The study was approved by the Departmental Ethics Committee and boards of Bahauddin Zakariya University, Multan, Pakistan as per the Declaration of Helsinki. Written informed consent was taken from all the participants who participated in this study.

Conflict of interest
No conflict of interest regarding the paper.

Received: 8 January 2024 / Accepted: 18 November 2024
Published online: 21 November 2024

References

1. Nie X, Liu X, Wang C, Wu Z, Sun Z, Su J, Yan R, Peng Y, Yang Y, Wang C, Cai S. Assessment of evidence on reported non-genetic risk factors of congenital heart defects: the updated umbrella review. BMC Pregnancy Childbirth. 2022;22(1):371.
2. Ali F, Ladak LA, Usmani AA, Raza HA, Siddiqui MT, Hasan B. Health-related quality of life in postcardiac interventional catheterization patients with congenital heart disease: a mixed-methods study protocol from Pakistan. BMJ Open. 2021;11(12):e052989.
3. Liu Y, Chen S, Zühlke L, Black GC, Choy MK, Li N, Keavney BD. Global birth prevalence of congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies. Int J Epidemiol. 2019;48(2):455–63.
4. Wu W, He J, Shao X. Incidence and mortality trend of congenital heart disease at the global, regional, and national level, 1990–2017. Medicine. 2020;99(23).
5. Shahid S, Akbar A. Conventional and non-conventional risk factors of cyanotic and acyanotic congenital heart diseases in Children of Southern Punjab, Pakistan. Pakistan Heart J. 2020;53(2).
6. Umer HM, Arshad MU, Mudassar F, Slam AU, Junaid S, Asghar RM, Aslam F. Pattern of congenital Heart diseases in Paediatric Age Group. J Rawalpindi Med Coll. 2018;22(1).
7. Rizvi SF, Mustafa G, Kundi A, Khan MA. Prevalence of congenital heart disease in rural communities of Pakistan. J Ayub Med Coll Abbottabad. 2015;27(1):124–7.

8. Pathan IH, Bangash SK, Khawaja AM. Spectrum of heart defects in children presenting for paediaric cardiac surgery. Pakistan Heart J. 2016;49(1).
9. Ekure EN, Bode-Thomas F, Sadoh WE, Orogade AA, Otaigbe BE, Ujunwa F, Sani UM, Asani M, Animasahun AB, Ogunkunle OC. Nigerian Pediatric Cardiology Study Group. Congenital heart defects in Nigerian children: Preliminary Data From the National Pediatric Cardiac Registry.
10. Kalisch-Smith JI, Ved N, Sparrow DB. Environmental risk factors for congenital heart disease. Cold Spring Harb Perspect Biol. 2020;12(3):a037234.
11. Liu S, Liu J, Tang J, Ji J, Chen J, Liu C. Environmental risk factors for congenital heart disease in the Shandong Peninsula, China: a hospital-based case–control study. J Epidemiol. 2009;19(3):122–30.
12. Haq FU, Jalil F, Hashmi S, Jumani MI, Imdad A, Jabeen M, Hashmi JT, Irfan FB, Imran M, Atiq M. Risk factors predisposing to congenital heart defects. Ann Pediatr Cardiol. 2011;4(2):117.
13. Van Der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, Roos-Hesselink JW. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. J Am Coll Cardiol. 2011;58(21):2241–7.
14. Mughal AR, Sadiq M, Hyder SN, Qureshi AU, Shah SS, Khan MA, Nasir JA. Socioeconomic status and impact of treatment on families of children with congenital heart disease. J Coll Physicians Surg Pak. 2011;21(7):398–402.
15. Feng Y, Yu D, Yang L, Da M, Wang Z, Lin Y, Ni B, Wang S, Mo X. Maternal lifestyle factors in pregnancy and congenital heart defects in offspring: review of the current evidence. Ital J Pediatr. 2014;40(1):1–7.
16. Lui GK, Fernandes S, McElhinney DB. Management of cardiovascular risk factors in adults with congenital heart disease. J Am Heart Association. 2014;3(6):e001076.
17. Yu D, Feng Y, Yang L, Da M, Fan C, Wang S, Mo X. Maternal socioeconomic status and the risk of congenital heart defects in offspring: a meta-analysis of 33 studies. PLoS ONE. 2014;9(10):e111056.
18. Shahid S, Khurram H, Billah B, Akbar A, Shehzad MA, Shabbir MF. Machine learning methods for predicting major types of rheumatic heart diseases in children of Southern Punjab, Pakistan. Front Cardiovasc Med. 2022;9:996225.
19. Mari MA, Cascudo MM, Alchieri JC. Congenital heart disease and impacts on child development. Brazilian J Cardiovasc Surg. 2016;31:31–7.
20. Pei L, Kang Y, Zhao Y, Yan H. Prevalence and risk factors of congenital heart defects among live births: a population-based cross-sectional survey in Shaanxi province, Northwestern China. BMC Pediatr. 2017;17(1):1–8.
21. Suresh KP, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. J Hum Reproductive Sci. 2012;5(1):7.
22. Dattalo P. Determining sample size: balancing power, precision, and practicality. Oxford University Press; 2008.
23. Shahid S, Khurram H, Lim A, Shabbir MF, Billah B. (2024). Prediction of cyanotic and acyanotic congenital heart disease using machine learning models. World J Clin Pediatr, *13*(4).
24. Shojaei B, Loripoor M, Sheikhfathollahi M, Aminzadeh F. (2021). The effect of walking during late pregnancy on the outcomes of labor and delivery: a randomized clinical trial. J Educ Health Promotion, *10*.
25. Höfer S, Benzer W, Alber H, Ruttmann E, Kopp M, Schussler G, Doering S. Determinants of health-related quality of life in coronary artery disease patients: a prospective study generating a structural equation model. Psychosomatics. 2005;46(3):212–23.
26. Garnier-Villarreal M, Jorgensen TD. Adapting fit indices for bayesian structural equation modeling: comparison to maximum likelihood. Psychol Methods. 2020;25(1):46.
27. Taasoobshirazi G, Wang S. The performance of the SRMR, RMSEA, CFI, and TLI: an examination of sample size, path size, and degrees of freedom. J Appl Quant Methods. 2016;11(3):31–9.
28. Kline RB. Principles and practice of structural equation modeling. Guilford; 2023.
29. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods Psychol Res Online. 2003;8(2):23–74.
30. Schubert AL, Hagemann D, Voss A, Bergmann K. Evaluating the model fit of diffusion models with the root mean square error of approximation. J Math Psychol. 2017;77:29–45.
31. Tennant A, Pallant JF. The root mean square error of approximation (RMSEA) as a supplementary statistic to determine fit to the Rasch model with large sample sizes. Rasch Meas Trans. 2012;25(4):1348–9.
32. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equation Modeling: Multidisciplinary J. 1999;6(1):1–55.
33. Kaplan D. Structural equation modeling: foundations and extensions. SAGE; 2008. Jul 23.
34. Chelo D, Nguefack F, Menanga AP, Um SN, Gody JC, Tatah SA, Ndombo PO. Spectrum of heart diseases in children: an echocardiographic study of 1,666 subjects in a pediatric hospital, Yaounde, Cameroon. Cardiovasc Diagnosis Therapy. 2016;6(1):10.
35. Otaigbe BE, Tabansi PN. Congenital heart disease in the Niger Delta region of Nigeria: a four-year prospective echocardiographic analysis: cardiovascular topic. Cardiovasc J Afr. 2014;25(6):265–8.
36. Du S, Feng Z, Wang W, Tian L, Wang Y. A structural equation model linking health literacy, self-efficacy and quality of life in adults with coronary heart disease. BMC Cardiovasc Disord. 2022;22(1):285.
37. Miao Q, Dunn S, Wen SW, Lougheed J, Reszel J, Lavin Venegas C, Walker M. Neighbourhood maternal socioeconomic status indicators and risk of congenital heart disease. BMC Pregnancy Childbirth. 2021;21:1–21.
38. Miao Q, Dunn S, Wen SW, Lougheed J, Maxwell C, Reszel J, Hafizi K, Walker M. Association of maternal socioeconomic status and race with risk of congenital heart disease: a population-based retrospective cohort study in Ontario, Canada. BMJ open. 2022;12(2):e051020.
39. Miao Q, Dunn S, Wen SW, Lougheed J, Sharif F, Walker M. Associations of congenital heart disease with deprivation index by rural-urban maternal residence: a population-based retrospective cohort study in Ontario, Canada. BMC Pediatr. 2022;22(1):476.

Publisher's note