Development and Internal Validation of a Clinical Scoring System for Early Detection of Congenital Heart Disease in Children

Mahrus Abdur Rahman[1]*, Prima Hari Nastiti[1], I Ketut Alit Utamayasa[1], Ayurveda Zaynabila Heriqbaldi[2] and Shabrina Nur Imanina[2]

[1]Division of Pediatric Cardiology, Department of Pediatrics, Faculty of Medicine, Universitas Airlangga /Dr. Soetomo General Hospital, Surabaya, Indonesia

[2]Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia

Received: 11 January 2026

Accepted: 28 October 2026

*Corresponding author: mahrus.a@fk.unair.ac.id

DOI 10.5001/omj.2026.86

Abstract

Objectives: Congenital heart disease (CHD) remains a leading cause of morbidity and mortality in children, particularly in low-resource settings. Early identification is essential to prevent life-threatening complications. This study aimed to develop and assess the diagnostic accuracy of a clinical scoring application for early detection of CHD in children.

Methods: A cross-sectional diagnostic study was conducted among 62 children referred for echocardiographic screening at Dr. Soetomo General Hospital, Surabaya from January to December 2025. The CHD Scoring Application generated a cumulative score based on 11 historical and clinical parameters, each weighted according to its diagnostic strength. Echocardiography was used as the diagnostic gold standard. Sensitivity, specificity, predictive values, likelihood ratios, ROC-AUC, Cohen's kappa, and bootstrap internal validation (B = 1000) were calculated.

Results: Among 62 subjects, 37 (59.7%) were confirmed with CHD. The mean score in CHD cases was significantly higher (6.76 ± 3.44) than in non-CHD (1.16 ± 0.90, $p < 0.001$). At a cutoff ≥ 3, the scoring system achieved a sensitivity of 75.7%, specificity of 96.0%, PPV of 96.6%, NPV of 72.7%, and overall accuracy of 83.9%. The AUC was 0.964 (95% CI 0.912–0.997), indicating excellent discriminative performance. Agreement with echocardiographic diagnosis was substantial (Cohen's $\kappa = 0.68$). Bootstrap internal validation demonstrated minimal optimism, confirming model robustness.

Conclusions: The CHD clinical scoring application demonstrated excellent diagnostic accuracy and internal validity, supporting its use as a feasible early detection tool for CHD in children, particularly in resource-limited healthcare settings where echocardiography access is restricted.

Keywords: Congenital Heart Disease; Pediatric Cardiology; Diagnostic Accuracy; Screening Tool; Validation Study; Clinical Scoring.

Introduction

Congenital heart disease (CHD) represents the most common congenital anomaly worldwide and remains a major contributor to infant morbidity and mortality. It encompasses a broad spectrum of structural abnormalities of the heart and great vessels, often resulting in impaired circulation, heart failure, or cyanosis. Classic epidemiologic data estimate the incidence of CHD at approximately 8–10 per 1,000 live births, with more recent population-based studies confirming a persistent and substantial burden, particularly in low- and middle-income countries.[1–3] In Indonesia, congenital heart anomalies remain the leading contributor to pediatric cardiovascular burden, with neonatal prevalence reaching 1,511.53 per 100,000 in 2023, and VSD, ASD, and PDA comprising the majority of lesions among children.[4,5]

In countries with limited resources, including Indonesia, CHD is frequently diagnosed late because of restricted access to echocardiography, a shortage of pediatric cardiologists, and uneven distribution of tertiary centers. These health-system constraints contribute to preventable morbidity and mortality from heart failure, pulmonary hypertension, and recurrent respiratory infections. National strategies to reduce under-five mortality have increasingly emphasized early recognition and timely referral of children with suspected CHD.[6]

Current methods for CHD detection include prenatal ultrasonography and postnatal pulse oximetry screening. While fetal echocardiography can identify many structural defects, its diagnostic performance and availability are limited by operator expertise, gestational age, and equipment capacity. Similarly, pulse oximetry screening has improved detection of critical cyanotic CHD, but its sensitivity is lower for acyanotic lesions.[7,8] As a result, a substantial proportion of children with CHD especially those with non-cyanotic lesions remain undetected and present later with complications.

Given these diagnostic gaps, there is a need for a practical, low-cost, and reliable clinical tool that can be used by healthcare workers at all levels, including primary and rural health facilities. A clinical scoring system that integrates easily identifiable signs and symptoms offers a promising approach for early CHD detection. In parallel, expansion of digital health technology enables mobile-based applications that can enhance accessibility, standardization, and speed of clinical decision-making.[9]

This study therefore aimed to develop and assess the diagnostic accuracy of a low-cost, mobile-based clinical scoring application for early detection of CHD in children.. By comparing the application's diagnostic performance with echocardiography as the gold standard, this study sought to provide evidence for its accuracy, feasibility, and potential integration into primary and secondary healthcare settings in resource-limited environments.

Methods

This analytical cross-sectional diagnostic study evaluated the validity and diagnostic accuracy of a newly developed clinical scoring application for early detection of CHD in children. The study was conducted at the Pediatric Cardiology Division, Dr. Soetomo General Hospital, Surabaya, Indonesia, a major tertiary referral center for pediatric cardiology in eastern Indonesia from January to December 2025. Ethical approval was granted by the Institutional Health Research Ethics Committee (No. 1360/KEPK/VII/2025), and written informed consent was obtained from parents or legal guardians.

Participants were children aged 0–18 years referred for echocardiography due to suspected cardiac abnormalities, including both inpatient and outpatient cases. Inclusion criteria were clinical manifestations suggestive of CHD (e.g., cyanosis, tachypnea, failure to thrive, recurrent respiratory infections, feeding difficulty, or abnormal cardiac auscultation). Exclusion criteria were previously diagnosed CHD, ongoing cardiac therapy or prior cardiac surgery, and incomplete data or refusal of consent. A total of 62 eligible consecutive participants were enrolled. Sample size calculation using a correlation formula ($\alpha = 0.01$, $\beta = 0.1$, $r = 0.5$) indicated a minimum requirement of 53 subjects.

The CHD Early Detection Application is a low-cost, mobile-based screening tool designed to support early recognition of CHD in primary and secondary healthcare settings. It consists of 11 clinical items encompassing patient history and physical findings, grouped into two categories:

1. Clinical history: prematurity, low birth weight (<2,500 g), cyanosis at birth, other congenital anomalies, feeding difficulty, and poor weight gain.

2. Physical examination findings: perioral cyanosis, tachypnea, tachycardia, chest retractions, and left chest bulging.

Each positive response was assigned a score of 1 or 2 according to its diagnostic weight, determined by the strength of its statistical association with CHD (Fisher's exact test and odds ratio) and physiological relevance. Items with OR $\geq$ 25 were assigned 2 points, while those with OR < 25 were assigned 1 point, reflecting a natural break in the OR distribution. Clinical scoring was performed by attending clinicians prior to echocardiographic

examination and without knowledge of the echocardiographic results, ensuring that index test assessors were blinded to the reference standard.

A cutoff score of ≥ 3 was pre-specified based on clinical judgment, requiring at least one high-weight symptom or a combination of lower-weight items. This threshold was subsequently confirmed using the Youden Index, which identified ≥ 2 as the statistically optimal cutoff; ≥ 3 was retained due to its superior specificity and PPV in the context of resource-limited referral triage. The application automatically calculated the total score and provided interpretation output in real-time.

Transthoracic two-dimensional (2D) echocardiography, performed by a pediatric cardiologist using a Philips EPIQ 7 ultrasound machine, served as the diagnostic gold standard. Echocardiographic assessment included measurement of cardiac chamber dimensions, septal integrity, valvular morphology, great vessel connections, and estimation of pulmonary artery pressure. All echocardiographic examinations were interpreted independently by two pediatric cardiologists blinded to the application results to minimize observer bias. A positive echocardiographic diagnosis was defined as the presence of any structural cardiac defect, including ventricular septal defect (VSD), atrial septal defect (ASD), patent ductus arteriosus (PDA), or complex cyanotic lesions. Structurally normal hearts were classified as CHD-negative.

All data were entered into a secured database and analyzed using IBM SPSS Statistics version 17.0 and Python (version 3.10). All participants had complete index test and reference standard results; no indeterminate or uninterpretable findings were recorded. Continuous variables were summarized as mean ± standard deviation (SD) or median with range, as appropriate. Categorical variables were expressed as frequencies and percentages. Group comparisons were performed using independent samples t-test or Mann-Whitney U test for continuous variables, and Fisher's exact test or chi-square test for categorical variables. Odds ratios (OR) with 95% confidence intervals (CI) were calculated for each scoring item using the Haldane-Anscombe correction where any cell count was zero.

Diagnostic performance was assessed using 2×2 contingency tables to calculate sensitivity, specificity, PPV, NPV, accuracy, and likelihood ratios, with 95% CIs derived using the Wilson score interval for proportions and the log-transformation method for likelihood ratios. Cohen's kappa (κ) with 95% CI was computed to assess agreement, with values of 0.61–0.80 interpreted as substantial and 0.81–1.00 as almost perfect agreement.[10] McNemar's test was performed to evaluate systematic discordance between the scoring system and echocardiographic diagnosis. ROC analysis was conducted to estimate the AUC with 95% CI derived using bootstrap resampling (B = 1000), and the optimal cutoff was determined using the Youden Index. Bootstrap internal validation (B = 1000) using Harrell's optimism-correction method was performed to assess overfitting; AUC values of 0.90–1.00 were considered excellent, 0.80–0.90 good, and 0.70–0.80 fair.[11,12] Internal consistency of the scoring system was assessed using Cronbach's alpha coefficient and item-total correlations, an alpha of ≥ 0.70 was considered acceptable.[13] A p-value < 0.05 was considered statistically significant.

Results

A total of 62 children were enrolled in this diagnostic-validation study, of whom 37 (59.7%) were confirmed to have CHD and 25 (40.3%) had structurally normal hearts as determined by echocardiography. Baseline characteristics of study participants are shown in Table 1.

Table 1: Baseline characteristics of study participants.

Characteristic	Category/Statistic	CHD (–) n=25	CHD (+) n=37	Total n=62	p-value
Age (months)	Mean ± SD (Min–Max)	58.54 ± 61.04 (2.5–204)	36.95 ± 51.43 (1–192)	45.65 ± 56.04 (1–204)	0.072
Sex, n (%)	Male	12 (40.0%)	18 (60.0%)	30 (100.0%)	1.000
	Female	13 (40.6%)	19 (59.4%)	32 (100.0%)	
Mode of Delivery, n (%)	Spontaneous vaginal	16 (45.7%)	19 (54.3%)	35 (100.0%)	0.435
	Cesarean section	9 (33.3%)	18 (66.7%)	27 (100.0%)	
Birth Weight (g)	Mean ± SD (Min–Max)	2814.00 ± 220.57 (2200–3300)	2637.84 ± 462.09 (1800–3900)	2708.87 ± 390.83 (1800–3900)	0.043
Nutritional Status, n (%)	Severe malnutrition	0 (0.0%)	2 (100.0%)	2 (100.0%)	<0.001
	Moderate malnutrition	0 (0.0%)	16 (100.0%)	16 (100.0%)	

	Mild malnutrition	16 (50.0%)	16 (50.0%)	32 (100.0%)	
	Normal	9 (75.0%)	3 (25.0%)	12 (100.0%)	
Echocardiographic Diagnosis	Ventricular Septal Defect (VSD)	—	17 (45.9%)	—	—
	Atrial Septal Defect (ASD)	—	7 (18.9%)	—	
	Patent Ductus Arteriosus (PDA)	—	4 (10.8%)	—	
	Complex lesions (e.g., combined VSD–PDA or TGA)	—	9 (24.3%)	—	

The mean age of participants was 45.65 ± 56.04 months. Children diagnosed with CHD tended to be younger than those without, though the difference did not reach statistical significance ($p = 0.072$). The younger age trend among CHD patients suggests that symptoms and clinical findings prompting referral often emerge early in life, particularly during infancy and toddler years when hemodynamic disturbances become clinically evident. Sex distribution was nearly equal, showing no significant association with CHD occurrence ($p = 1.000$). Regarding mode of delivery, 35 children were born by spontaneous vaginal delivery. Although the CHD proportion was slightly higher among caesarean births (66.7%) than spontaneous deliveries, this difference was not statistically significant ($p = 0.435$). The higher rate in the caesarean group might reflect a tendency toward surgical delivery among high-risk pregnancies, including those with fetal cardiac anomalies detected or suspected prenatally. The mean birth weight of the cohort was 2708.87 ± 390.83 g, with a significant difference between CHD and non-CHD groups (2637.84 ± 462.09 g vs 2814.00 ± 220.57 g, $p = 0.043$). This finding indicates that low birth weight (LBW) remains a significant perinatal correlate of congenital cardiac anomalies. Fetal hypoxia, maternal infection, or placental insufficiency may contribute to this association, consistent with the multifactorial origins of CHD.

Marked differences were observed in nutritional status between CHD and non-CHD children ($p < 0.001$). Among CHD patients, 48.6% Suffered from moderate-to-severe malnutrition, while most non-CHD children were well nourished or mildly undernourished. These results demonstrate a clear association between chronic cardiac disease and growth failure, likely mediated by elevated metabolic demand, inadequate caloric intake due to early fatigue during feeding, and malabsorption secondary to venous congestion. Among CHD+ patients, the most common diagnoses were VSD (45.9%).

Table 2 presents the association of each clinical item with CHD status. Eight of eleven items demonstrated statistically significant associations with CHD ($p < 0.05$). Five items: feeding difficulty, poor weight gain, perioral cyanosis, tachypnea, and tachycardia, showed the strongest associations (OR range 25.00–92.56; all $p < 0.001$) and were assigned a score of +2. Six items: preterm birth, low birth weight, cyanosis at birth, other congenital anomalies, chest retraction, and left chest bulging, were assigned a score of +1. A natural break in OR distribution (OR 17 vs 25) justified this two-tier weighting scheme. Although other congenital anomalies showed an OR < 1 (OR = 0.26), it was retained in the scoring system based on clinical judgment. OR values for items with zero cell counts were estimated using the Haldane-Anscombe correction. The total possible score ranges from 0 to 16. The mean total score among CHD-positive patients was 6.76 ± 3.44, compared with 0.48 ± 0.77 among non-CHD patients ($p < 0.001$).

Table 2: Bivariate analysis – association of each clinical item with CHD status.

Criterion	CHD (–) *n (%)*	CHD (+) *n (%)*	OR (95% CI)	p-value	Points
Preterm birth	1 (4.0%)	8 (21.6%)	6.62 (0.77–56.73)	0.072	+1
Low birth weight	1 (4.0%)	13 (35.1%)	13.00 (1.57–107.36)	0.005	+1
Cyanotic appearance at birth	0 (0.0%)	9 (24.3%)	17.00 (0.94–306.99)	0.008	+1
Other congenital anomalies*	8 (32.0%)	4 (10.8%)	0.26 (0.07–0.98)	0.052	+1
Chest retraction (suprasternal/ intercostal/subcostal)	0 (0.0%)	9 (24.3%)	17.00 (0.94–306.99)	0.008	+1
Left chest bulging	0 (0.0%)	5 (13.5%)	8.63 (0.46–163.45)	0.076	+1
Feeding difficulty	0 (0.0%)	16 (43.2%)	39.14 (2.22–691.32)	<0.001	+2
Poor weight gain	1 (4.0%)	27 (73.0%)	64.80 (7.72–544.16)	<0.001	+2
Bluish lips/cyanosis	0 (0.0%)	12 (32.4%)	25.00 (1.40–445.16)	<0.001	+2
Tachypnea	0 (0.0%)	22 (59.5%)	74.03 (4.19–1309.24)	<0.001	+2
Tachycardia	0 (0.0%)	24 (64.9%)	92.56 (5.21–1643.32)	<0.001	+2

OR = odds ratio; CI = confidence interval. Fisher's exact test used for all comparisons. OR calculated with Haldane-Anscombe correction (+0.5) where any cell count = 0. *Other congenital anomalies retained in the scoring system based on clinical judgment despite OR < 1.

The distribution of all clinical history and physical findings by CHD status is presented in Table 3. Percentages are calculated as column percentages (proportion within each diagnostic group). Statistically significant differences were observed in 9 of 11 items ($p < 0.05$). Total score differed markedly between groups: mean 6.76 ± 3.44 (median 8, range 0–12) in CHD+ versus 0.48 ± 0.77 (median 0, range 0–3) in CHD− ($p < 0.001$). At the selected cutoff of ≥ 3, 28 of 37 CHD+ patients (75.7%) were classified as likely CHD, compared to only 1 of 25 CHD− patients (4.0%).

Table 3: Distribution of clinical history and physical findings by CHD status.

Criterion	CHD (–) n (%)	CHD (+) n (%)	p-value	Score
History				
Preterm birth	1 (4.0%)	8 (21.6%)	0.072	+1
Low birth weight	1 (4.0%)	13 (35.1%)	0.005	+1
Cyanotic appearance at birth	0 (0.0%)	9 (24.3%)	0.008	+1
Other congenital anomalies*	8 (32.0%)	4 (10.8%)	0.052	+1
Feeding difficulty	0 (0.0%)	16 (43.2%)	<0.001	+2
Poor weight gain	1 (4.0%)	27 (73.0%)	<0.001	+2
Physical Appearance				
Bluish lips/perioral cyanosis	0 (0.0%)	12 (32.4%)	<0.001	+2
Tachypnea	0 (0.0%)	22 (59.5%)	<0.001	+2
Tachycardia	0 (0.0%)	24 (64.9%)	<0.001	+2
Chest retraction	0 (0.0%)	9 (24.3%)	0.008	+1
Left chest bulging	0 (0.0%)	5 (13.5%)	0.076	+1
Total Score				
Mean ± SD (Median, Range)	0.48 ± 0.77 (0, 0–3)	6.76 ± 3.44 (8, 0–12)	<0.001	—
Likely CHD (score ≥ 3)	1 (4.0%)	28 (75.7%)	<0.001	—
Not CHD (score < 3)	24 (96.0%)	9 (24.3%)	—	—

Data presented as n (column %). p-values from Fisher's exact test. Score column indicates the point value assigned to each item. *Other congenital anomalies retained based on clinical judgment (OR < 1 in this sample).

The diagnostic performance calculations for the cut-off score ≥ 3 are presented in Table 4. At the selected cutoff score of ≥ 3, the scoring system demonstrated a sensitivity of 0.757 (95% CI 0.599–0.866), specificity of 0.960 (95% CI 0.805–0.993), positive predictive value (PPV) of 0.966 (95% CI 0.828–0.994), and negative predictive value (NPV) of 0.727 (95% CI 0.558–0.849). Overall accuracy was 0.839 (95% CI 0.728–0.910). The positive likelihood ratio (LR+) was 18.92 (95% CI 2.75–130.22), indicating strong confirmatory value, while the negative likelihood ratio (LR−) was 0.253 (95% CI 0.143–0.450). Inter-rater agreement, measured by Cohen's kappa, was 0.681 (95% CI 0.501–0.862), indicating substantial agreement. The McNemar test was statistically significant ($p = 0.021$), reflecting that discordance was predominantly driven by false-negatives. ROC curve analysis yielded an AUC of 0.964 (95% CI 0.912–0.997), indicating excellent discriminatory ability. The Youden Index identified a statistically optimal cutoff of ≥ 2 (sensitivity 94.6%, specificity 92.0%, J = 0.866). However, a cutoff of ≥ 3 was selected for clinical use, yielding a sensitivity of 75.7%, specificity of 96.0%, PPV of 96.6%, and NPV of 72.7%, as further discussed below.

Table 4: Diagnostic performance of the CHD early detection scoring system (Cutoff ≥ 3).

Parameter	Estimate	95 % CI	Interpretation
Sensitivity	0.757	0.599–0.866	Detects 75.7% of CHD cases
Specificity	0.960	0.805–0.993	Excludes 96.0% of non-CHD
PPV	0.966	0.828–0.994	Positive result almost always true
NPV	0.727	0.558–0.849	Some missed cases among negatives
Accuracy	0.839	0.728–0.910	Correctly classifies 83.9% of patients
LR+	18.92	2.75–130.22	Strong confirmatory evidence
LR–	0.253	0.143–0.450	Moderate rule-out value
Cohen’s κ	0.681	0.501–0.862	Substantial agreement
McNemar p	0.021	—	Discordance driven by false-negatives
AUC	0.964	0.912–0.997	Excellent discrimination

95% CI for sensitivity, specificity, PPV, NPV, and accuracy calculated using Wilson score interval. 95% CI for LR+ and LR− calculated using log-transformation method. 95% CI for Cohen's κ calculated using Fleiss standard error formula. 95% CI for AUC calculated using bootstrap method (B = 1000). PPV = positive predictive value; NPV = negative predictive value; LR = likelihood ratio; AUC = area under the curve.

Cronbach's $\alpha = 0.766$, indicating acceptable internal consistency. Item-total correlations ranged from −0.19 to 0.82. Three items (preterm birth, low birth weight, and other congenital anomalies) showed negative item-total correlations, suggesting limited contribution to scale coherence and warranting further evaluation.

Bootstrap internal validation was performed using Harrell's optimism-correction method (B = 1000 resamples) to assess potential overfitting of the scoring system. Results are presented in Table 5. Optimism estimates were negligible across all performance metrics (range: −0.005 to +0.001), indicating that the apparent performance measures were reliable and not meaningfully inflated by overfitting. This finding is expected given that the scoring system uses fixed integer weights based on clinical categories rather than optimised regression coefficients, thereby inherently limiting the risk of data-driven overfitting. Beyond discrimination and calibration, internal consistency was evaluated to assess the coherence of the scoring items. Cronbach's alpha was 0.766 (acceptable), with item-total correlations ranging from 0.18 to 0.62, suggesting that all eleven items collectively and consistently measure the underlying likelihood of CHD.

Table 5: Bootstrap internal validation (B = 1000, Harrell's Optimism-Correction Method).

Parameter	Apparent	Optimism	Corrected	95% CI (Bootstrap)
Sensitivity	0.757	−0.005	0.761	0.595–0.886
Specificity	0.960	+0.001	0.959	0.862–1.000
PPV	0.966	−0.000	0.966	0.880–1.000
NPV	0.727	−0.001	0.728	0.555–0.871
Accuracy	0.839	−0.002	0.841	0.742–0.919
AUC	0.964	−0.001	0.965	0.909–0.997

Optimism = mean difference between bootstrap sample performance and original sample performance across 1000 bootstrap resamples. Corrected = apparent − optimism. 95% CI derived from percentile bootstrap method. AUC = area under the ROC curve.

Discussion

The clinical scoring system developed in this study comprises 11 parameters derived from statistically significant variables differentiating CHD from non-CHD patients. These parameters reflect three diagnostic domains: perinatal risk factors, feeding and growth indicators, and cardiorespiratory signs. The scoring model achieved a high diagnostic accuracy, confirming that these simple bedside findings can collectively identify children with CHD with excellent discrimination and specificity. This aligns with prior work demonstrating that combinations of clinical features can be used as low-cost triage tools to select children who require echocardiographic confirmation.[8,14] While individual clinical indicators included in this scoring system (tachypnea, tachycardia, cyanosis, feeding difficulty , low birth weight) are individually recognized in the literature, the novelty and contribution of this work lies in three distinct areas: (1) systematic quantitative integration of weighted clinical parameters calibrated to an Indonesian pediatric population and validated against echocardiography; (2) implementation on a mobile-based platform enabling automated, real-time, standardized scoring, addressing the inter-observer variability inherent in informal clinical assessment; and (3) local derivation and internal evaluation, which is essential because diagnostic tools developed in epidemiologically different settings may not perform equivalently in Indonesia's health system context. The scoring tool is intended to supplement, not replace, clinical judgement; any child with a single high-risk feature (central cyanosis, severe tachycardia, refractory respiratory distress) warrants immediate echocardiographic referral regardless of total score. Below, each variable is discussed in detail.

Preterm infants have an increased risk of both structural and functional cardiac abnormalities. Incomplete septation and delayed ductal closure are common in infants born before 37 weeks, predisposing them to patent ductus arteriosus (PDA) and secondary pulmonary hypertension. Previous studies have shown that approximately 14% of preterm infants have a persistent patent ductus arteriosus, highlighting the higher vulnerability of premature neonates to persistent shunting lesions.[15] Prematurity also increases vulnerability to hypoxia and infection, compounding the hemodynamic stress on an immature myocardium. Similarly, low birth weight (LBW) was significantly associated with CHD in this study (p = 0.005). LBW infants have a higher risk of congenital malformations due to intrauterine hypoxia, maternal anaemia, or nutritional deficiency affecting cardiac morphogenesis,[16,17] and may also reflect placental insufficiency or maternal illness during organogenesis. Cyanotic appearance at birth suggests right-to-left shunting or parallel circulation, as seen in Tetralogy of Fallot, transposition of the great arteries, or pulmonary atresia. Unlike respiratory causes of desaturation, cyanosis in CHD is typically persistent and unresponsive to supplemental oxygen. Among 524 NICU admissions, 50 neonates (9.5%) diagnosed with critical CHD all presented with central cyanosis at admission,[18] consistent with meta-analytic data supporting the high predictive value of hypoxaemia for critical CHD.[7] Beyond isolated perinatal factors, the coexistence of non-cardiac anomalies increases the likelihood of CHD through shared developmental pathways involving transcription factors such as NKX2-5, GATA4, and TBX5.[19] Although this variable did not reach strong statistical significance in the present sample, its inclusion improves sensitivity for syndromic CHD

detection, supported by a pooled analysis of 102 studies reporting a CHD prevalence of 49.9% among individuals with Down syndrome.[20]

Feeding difficulty and failure to thrive were among the strongest predictors of CHD in this cohort, present in 43.2% and 73.0% of CHD patients, respectively (both $p < 0.001$). These findings are consistent with prior studies showing that hemodynamically significant shunt lesions increase metabolic demand while simultaneously impairing oral intake because of fatigue, tachypnea, and dyspnea during feeding.[21,22] Feeding difficulty arises from fatigue, dyspnea, and poor perfusion during feeding, which increases oxygen demand while reducing intake, a functional marker of heart failure particularly associated with left-to-right shunt lesions. A previous review of 1,107 participants reported a pooled prevalence of feeding and swallowing difficulties of 42.9%, with 32.9% demonstrating aspiration.[23] Failure to thrive results from the combined effects of chronic hypoxemia, elevated metabolic rate, and recurrent infections; a previous study reported underweight, stunting, and wasting in 54.2%, 43.4%, and 53% of 83 pediatric CHD patients, respectively.[24,25] Nutritional rehabilitation has been shown to improve both postoperative and long-term outcomes,[24,26,27] underscoring that growth faltering is not merely a consequence of CHD but a sensitive indirect marker of chronic cardiovascular compromise warranting early identification.

Cyanosis localized to the lips and tongue signifies arterial desaturation from right-to-left shunt or mixing lesions. Its absence in all non-CHD children and presence in 32.4% of CHD patients ($p < 0.001$) highlight its high specificity. Central cyanosis becomes clinically detectable when deoxygenated haemoglobin reaches approximately 3–5 g/dL and, unlike acrocyanosis, persists despite warming or supplemental oxygen, a distinction particularly important in primary-care settings lacking oximetry.[28,29] Tachypnea, present in nearly 60% of CHD cases ($p < 0.001$), arises from pulmonary congestion secondary to left-to-right shunt lesions and should prompt suspicion of heart failure when prolonged without fever or wheezing.[30,31] Persistent tachycardia, noted in 64.9% of CHD patients ($p < 0.001$), arises as a compensatory response to reduced stroke volume, with higher Ross scores correlating with worse ventricular function and poorer outcomes.[32,33] Intercostal or subcostal retractions (24.3%, $p = 0.008$), typically in those with large VSDs or PDAs, reflect increased work of breathing from elevated pulmonary venous pressures and serve as accessible bedside indicators of heart failure severity.[34,35] Visible precordial bulging, though less frequent (13.5%), develops from chronic right-ventricular dilation; imaging studies have shown precordial asymmetry correlates with right-ventricular hypertrophy on echocardiography,[36,37] and a study of 50 children with CHD reported bulging in 26% of cases.[38] These two signs add specificity to the score by identifying children with advanced or uncorrected disease.

The variables in the CHD Early Detection score represent a coherent pathway linking structural defects to hemodynamic changes and observable clinical signs. The strong clustering of major symptoms explains the high diagnostic power of the scoring system. When several high-weight features co-occur, the PPV of the scoring system reaches 96.6%, indicating that a positive score almost certainly confirms CHD. Compared with stand-alone pulse oximetry screening, which primarily detects duct-dependent cyanotic lesions, the present score captures both cyanotic and acyanotic, failure-related lesions, making it more comprehensive for use in primary care and low-resource contexts.[7,8,14]

In this cohort, echocardiography confirmed CHD in 59.7% of screened children, with with VSD being the most frequent lesion, followed by complex cyanotic defects, ASD, and PDA. This distribution mirrors global epidemiologic patterns in which VSD accounts for roughly 30–40% of CHD cases.[2,3] In many low- and middle-income settings, late presentation leads to complications such as pulmonary hypertension and Eisenmenger physiology. Integrating a structured clinical score into routine pediatric assessment could support earlier identification and referral, reducing the proportion of children who reach tertiary centers only when advanced complications have developed. Such an approach aligns with national health priorities and global targets for reducing preventable under-five mortality.[6]

This study demonstrates that the clinical scoring system, implemented both in manual form and as a mobile application, provides excellent diagnostic performance for early detection of CHD. The AUC of 0.964 (95% CI 0.912–0.997) and specificity of 96% indicate a strong ability to differentiate CHD from non-CHD cases, exceeding or matching the performance of several previously published screening tools such as the New Delhi score and other simplified algorithms.[6,8] These findings support the scoring system as an accurate and practical triage tool in low-resource settings where echocardiography availability remains limited.[1,3] Bootstrap internal validation

demonstrated minimal optimism across all performance metrics, suggesting that the scoring system is robust and unlikely to be substantially affected by overfitting despite the relatively modest sample size.

The high specificity (96%) and PPV (96.6%) indicate that a positive score strongly predicts true CHD. Although sensitivity (75.7%) and NPV (72.7%) are more modest, missed cases are likely to involve smaller, hemodynamically less significant defects that may not be immediately life-threatening. Although the Youden Index identified a cutoff of ≥ 2 as statistically optimal ($J = 0.866$, sensitivity 94.6%, specificity 92.0%), a cutoff of ≥ 3 was selected based on clinical considerations. In the context of confirmatory diagnostic triage — rather than population-based screening — higher specificity and PPV are preferable to minimize unnecessary echocardiographic referrals in resource-limited settings. Furthermore, a score of ≥ 2 can be reached by a single item with weight 2 alone, which is clinically insufficient as a basis for referral, whereas ≥ 3 requires a meaningful combination of symptoms. The substantial agreement between the scoring system and echocardiographic diagnosis (Cohen's $\kappa = 0.68$) further supports its clinical reliability.

A notable strength of this study is the evaluation of the scoring system in a mobile-based format. Digital health applications can standardize data entry, automate score calculation, and provide instant interpretation, thereby reducing inter-observer variability and supporting frontline health workers in remote or underserved areas.[7] In Indonesia, where access to pediatric echocardiography and subspecialty care is uneven, this technology has the potential to enhance early recognition and reduce late presentations that contribute to pulmonary hypertension and advanced heart failure.[6]

This study has several limitations. The small single-centre sample ($n = 62$) limits statistical precision and generalisability, as reflected in the wide confidence interval for LR+. Furthermore, participants were recruited from a tertiary referral centre among children already suspected of CHD, resulting in a high disease prevalence (59.7%) that may not reflect the broader population of children presenting to primary care. This referral-based selection introduces spectrum bias, which likely inflates PPV and may limit the applicability of the reported diagnostic performance to lower-prevalence settings. As a single-cohort derivation study, the item weights and cutoff were derived from the same dataset used for performance evaluation; although bootstrap internal validation confirmed negligible optimism, this approach cannot fully substitute for independent external validation. External validation in an independent, multicentre cohort remains necessary before broader clinical implementation.

The selection of cutoff ≥ 3, while clinically justified, yields a sensitivity of 75.7%, meaning approximately one in four CHD cases may be missed at the screening level. Additionally, three scoring items showed negative item-total correlations, and one item (other congenital anomalies) was retained based on clinical judgment despite $OR < 1$, which may introduce subjectivity into the model. Finally, symptom-based scoring is inherently limited in detecting haemodynamically insignificant defects that present without overt clinical features. Future studies should prioritise external multicentre validation with larger and more diverse samples, data-driven refinement of item weights to address poor item-total correlations, and implementation research evaluating the feasibility and cost-effectiveness of the mobile-based scoring application within national child health programmes.

Conclusion

The clinical scoring system evaluated in this study demonstrates high diagnostic accuracy (accuracy 83.9%, AUC 0.964), excellent specificity (96.0%), and acceptable internal consistency (Cronbach's $\alpha = 0.766$), confirming its value as an early CHD detection tool in children, particularly in resource-limited healthcare environments. Because it relies on simple clinical signs, perinatal history, and a mobile-application platform, the tool can be integrated into primary-care and community-level screening programs without substantial additional resources. Wider implementation has the potential to improve early CHD recognition, optimize referral pathways, and reduce morbidity and mortality associated with delayed diagnosis. Further multicenter validation and implementation research are recommended to support broader national adoption and integration into child health policies.

Disclosure

The authors declared no conflicts of interest. No funding was received for this study.

References

1. Dolk H, McCullough N, Callaghan S, Casey F, Craig B, Given J, et al. Risk factors for congenital heart disease: The Baby Hearts Study, a population-based case-control study. PLoS One. 2020;15(2):e0227908. doi:10.1371/journal.pone.0227908 PubMed PMID: 32092068; PubMed Central PMCID: PMC7039413.

2. Hoffman JI, Kaplan S. The incidence of congenital heart disease. Journal of the American college of cardiology. 2002;39(12):1890–900.

3. Van Der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. Journal of the American College of Cardiology. 2011;58(21):2241–7.

4. Iqhrammullah M, Rampengan DD, Amri I, Ramadhan RN, Ghifari MH, Khansa F, et al. Geographical inequalities and temporal trends in pediatric cardiovascular diseases in Indonesia: a 34-year global burden of disease analysis. Frontiers in Public Health. 2025;13:1688700.

5. Ismail MT, Hidayati F, Krisdinarti L, Noormanto N, Nugroho S, Wahab AS. Epidemiological profile of congenital heart disease in a national referral hospital. ACI (Acta Cardiologia Indonesiana). 2015;1(2).

6. Kemenkes RI. Pedoman Penatalaksanaan Gagal Jantung Anak di Fasilitas Pra-Rujukan. Jakarta: Kementerian Kesehatan Republik Indonesia; 2024.

7. Thangaratinam S, Brown K, Zamora J, Khan KS, Ewer AK. Pulse oximetry screening for critical congenital heart defects in asymptomatic newborn babies: a systematic review and meta-analysis. The Lancet. 2012;379(9835):2459–64.

8. Sands A, Craig B, Mulholland C, Patterson C, Dornan J, Casey F. Echocardiographic screening for congenital heart disease: a randomized study. J Perinat Med. 2002;30(4):307–12. doi:10.1515/JPM.2002.045 PubMed PMID: 12235719.

9. Majani N, Chillo P, Slieker MG, Sharau G, Mlawi V, Mongella S, et al. Newborn Screening for Critical Congenital Heart Disease in a Low-Resource Setting; Research Protocol and Preliminary Results of the Tanzania Pulse Oximetry Study. Global Heart. 2022 May 26;17(1). doi:10.5334/gh.1110

10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. biometrics. 1977;159–74.

11. Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988;240(4857):1285–93.

12. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.

13. Cronbach LJ. Coefficient alpha and the internal structure of tests. psychometrika. 1951;16(3):297–334.

14. Izhar FM, Abqari S, Shahab T, Ali SM. Clinical score to detect congenital heart defects: Concept of second screening. Ann Pediatr Cardiol. 2020;13(4):281–8. doi:10.4103/apc.APC_113_19 PubMed PMID: 33311915; PubMed Central PMCID: PMC7727906.

15. Deselina B, Putra ST, Suradi R. Prevalence of patent ductus arteriosus in premature infants at the Neonatal Ward, Cipto Mangunkusumo Hospital, Jakarta. Paediatrica Indonesiana. 2004;44(6):223–7. doi:10.14238/pi44.6.2004.223-7

16. Damayantie V, Rahayuningsih SE, Afriandi I. Congenital Heart Disease Characteristics in Low Birth Weight Infants at Dr. Hasan Sadikin General Hospital in 2010–2014. Althea Medical Journal. 2019 Oct 18;6(3):115–22. doi:10.15850/amj.v6n3.1533

17. Chung SH, Kim CY, Lee BS, and the Korean Neonatal Network. Congenital Anomalies in Very-Low-Birth-Weight Infants: A Nationwide Cohort Study. Neonatology. 2020;117(5):584–91. doi:10.1159/000509117 PubMed PMID: 32772029.

18. Abou-Taleb A, Abdelhamid MA, Bahkeet MAM. Clinical profile of cyanotic congenital heart disease in neonatal intensive care unit at Sohag University Hospital, Upper Egypt. Egyptian Journal of Medical Human Genetics. 2017 Jan 1;18(1):47–51. doi:10.1016/j.ejmhg.2016.01.003

19. Blue GM, Kirk EP, Giannoulatou E, Sholler GF, Dunwoodie SL, Harvey RP, et al. Advances in the Genetics of Congenital Heart Disease: A Clinician′s Guide. J Am Coll Cardiol. 2017 Feb 21;69(7):859–70. doi:10.1016/j.jacc.2016.11.060 PubMed PMID: 28209227.

20. Xu CX, Chen L, Cheng Y, Du Y. Prevalence of congenital heart defects in people with Down syndrome: a systematic review and meta-analysis. J Epidemiol Community Health. 2025 May 9;79(6):445–50. doi:10.1136/jech-2023-220638 PubMed PMID: 39805637.

21. Ravishankar C. Feeding challenges in the newborn with congenital heart disease. Curr Opin Pediatr. 2022 Oct 1;34(5):463–70. doi:10.1097/MOP.0000000000001162 PubMed PMID: 36000379.

22. Menon G, Poskitt E. Why does congenital heart disease cause failure to thrive? Archives of disease in childhood. 1985;60(12):1134–9.

23. Norman V, Zühlke L, Murray K, Morrow B. Prevalence of Feeding and Swallowing Disorders in Congenital Heart Disease: A Scoping Review. Front Pediatr. 2022 Apr 5;10:843023. doi:10.3389/fped.2022.843023 PubMed PMID: 35450100; PubMed Central PMCID: PMC9016225.

24. Lee J, Marshall T, Buck H, Pamela M, Daack-Hirsch S. Growth Failure in Children with Congenital Heart Disease. Children. 2025;12(5):616.

25. Flynn J, Utamayasa IKA, Rahman MA, Widjaja NA. Risk factors of malnutrition and growth faltering in children with congenital heart disease. J Med Pharm Chem Res. 2025;7(5):1026–37.

26. Mitchell I, Logan R, Pollock J, Jamieson M. Nutritional status of children with congenital heart disease. Heart. 1995;73(3):277–83.

27. Argent AC, Balachandran R, Vaidyanathan B, Khan A, Kumar RK. Management of undernutrition and failure to thrive in children with congenital heart disease in low- and middle-income countries. Cardiology in the Young. 2017 Dec;27(S6):S22–30. doi:10.1017/S104795111700258X

28. Martin L, Khalil H. How much reduced hemoglobin is necessary to generate central cyanosis? Chest. 1990;97(1):182–5.

29. Park MK, Salamat M. Park′s Pediatric Cardiology for Practitioners [Internet]. Seventh Edition. Philadelphia: Elsevier; 2020 [cited 2026 Apr 8]. Available from: https://shop.elsevier.com/books/parks-pediatric-cardiology-for-practitioners/park/978-0-323-68107-0

30. Levene R, Pollak-Christian E, Garg A, Keenaghan M. It Is Not Always Sepsis: Fatal Tachypnea in a Newborn. Case Rep Pediatr. 2018 Feb 21;2018:7858192. doi:10.1155/2018/7858192 PubMed PMID: 29527378; PubMed Central PMCID: PMC5840949.

31. Masarone D, Valente F, Rubino M, Vastarella R, Gravino R, Rea A, et al. Pediatric Heart Failure: A Practical Guide to Diagnosis and Management. Pediatrics & Neonatology. 2017 Aug 1;58(4):303–12. doi:10.1016/j.pedneo.2017.01.001

32. Hinton RB, Ware SM. Heart Failure in Pediatric Patients with Congenital Heart Disease. Circ Res. 2017 Mar 17;120(6):978–94. doi:10.1161/CIRCRESAHA.116.308996 PubMed PMID: 28302743; PubMed Central PMCID: PMC5391045.

33. Ross RD. The Ross classification for heart failure in children after 25 years: a review and an age-stratified revision. Pediatric cardiology. 2012;33(8):1295–300.

34. UTMB Health Pediatrics. Heart Failure [Internet]. 2020 [cited 2025 Dec 9]. Available from: https://www.utmb.edu/pedi_ed/CoreV2/Cardiology/Cardiology13.html

35. Maduabuchi CJ, Uleanya ND, Onukwuli VO, Chinawa AT, Obidike EO, Chukwu BF. Observed causes of severe respiratory distress among children with congenital heart disease. Nigerian Journal of Medicine. 2020;29(4):655–9.

36. Mertens LL, Friedberg MK. Imaging the right ventricle—current state of the art. Nature Reviews Cardiology. 2010;7(10):551–63.

37. Jone PN, Kim JS, Burkett D, Jacobsen R, Alvensleben JV. Cardiovascular Diseases. In: Bunik M, Hay WW, Levin MJ, Abzug MJ, editors. Current Diagnosis & Treatment: Pediatrics, 26e [Internet]. New York, NY: McGraw-Hill Education; 2022 [cited 2025 Dec 9]. Available from: accesspediatrics.mhmedical.com/content.aspx?aid=1190364208

38. G S, Kulkarni M, Kawade RV. Spectrum of Congenital heart diseases in children at a tertiary care hospital of Western Maharashtra- a cross-sectional study. European Journal of Cardiovascular Medicine [Internet]. 2024 Sep 24 [cited 2025 Dec 9];14:262–8.