Pediatric Rheumatology

RESEARCH ARTICLE Open Access

Novel multiple Z-score models for detection of coronary artery dilation: application in Kawasaki disease

Ho-Chang Kuo[1,2,3,4], Shih-Hsin Chen[5*], I-Fei Chen[6], Wen-Ing Cheng[6*], Shih-Feng Liu[4,7], Mindy Ming-Huey Guo[1,2,3], Yu-Chi Lin[3] and Yi-Hui Chen[3,8*]

Abstract

Background This study aims to develop Z-Score models to normalize measurements of three coronary arteries and enhance the diagnosis of Kawasaki disease (KD) in children from newborns to 10 years old. Developing a reliable Z-Score model is challenging, as some existing models fail the normality test. Overcoming these challenges is crucial for improving KD diagnosis.

Method Detailed measurements of the left main coronary artery (LCA), left anterior descending coronary artery (LAD), and right coronary artery (RCA) were collected, along with patient demographics such as age, height, weight, and body surface area (BSA). Several Z-Score models, named the Kuo Z-Score models, were proposed, with separate designs for different coronary arteries and different age groups, resulting in multiple Z-Score models. The Z-Score model for the RCA employs the Box-Cox method for data transformation. Finally, we tested various age group combinations, selecting models that passed the Anderson–Darling normality test and had higher R-square values for robustness and best data fit.

Results The study included 1180 participants free from coronary or heart diseases. The Kuo Z-Score models were optimized for LCA, LAD, and RCA across the five age groups 0–6 years, 6–7 years, 7–8 years, 8–9 years, and 9–10 years. Only the normality test for the RCA in the 7–8 year age group failed. The proposed model fitted to the normality assumption outperforming the other models.

Conclusion The Kuo Z-Score models, applicable across a broad age range, provides robust identification of coronary artery dilatation and aneurysm in KD. The models' capability to normalize diverse data sets marks a significant advancement in KD diagnostic sensitivity, aiding in better clinical decision-making and potentially improving patient outcomes.

Keywords Kawasaki disease, Z-Score, Normality Test, Box-Cox Transformation

*Correspondence:
Shih-Hsin Chen
shchen@mail.tku.edu.tw
Wen-Ing Cheng
yaym91@gmail.com
Yi-Hui Chen
cyh@gap.cgu.edu.tw
Full list of author information is available at the end of the article

Background

Kawasaki disease (KD), also known as mucocutaneous lymph node syndrome, is a systemic vasculitis that affects small to medium-sized blood vessels, including the coronary arteries [1]. KD predominantly occurs in children aged 6 months to 5 years, with an incidence rate of about 0.06% whereas the incidence rate was 17.6 for 100,000 children under 5 years. In Taiwan, as per the National Health Insurance Administration, there are 800 to 1000 new KD cases reported yearly, mostly in children under 5 years, who constitute 94% of all cases. Even though not every patient experienced coronary artery lesions (CALs) [2], 35% of KD patients still have this symptom [3]. The overall mortality risk associated with KD is approximately 0.17%, underscoring the importance of early diagnosis.

The primary clinical manifestations of KD include coronary artery aneurysm [4]. Given the variation in coronary artery diameter with age, weight, and height, objective measurements are essential for the early detection of these conditions. The Japanese criteria [5] specify that an arterial diameter greater than 3.0 mm is indica- tive of KD CALs in children under 5 years. However, there is a notable variation in coronary artery diameters.

Consequently, Z-Scores, which adjust coronary artery diameter data based on body surface area (BSA) with mean and standard deviation, are employed for more accurate detection [6–11]. Olivieri et al. [9]. optimized linear regression models, incorporating quadratic or cubic BSA terms, while McCrindle et al. [8] used BSA-based nonlinear regression to track echocardiographic changes in Kawasaki Disease (KD) patients. Dallaire and Dahdah [6] explored multiple regression func- tions, and Lin et al. [7] applied regression analysis to a Taiwanese cohort, high- lighting the influence of sample size on model accuracy. Recent studies emphasize the importance of selecting the appropriate Z-Score system for diagnosing coronary abnormalities in KD, with Yoo [12], Kim [13], and Lorenzoni et al. [14]. underscoring the variability and clinical impact of different Z-Score formulas.

In our research, we employed the Anderson–Darling test to assess the normality of Z-Scores derived from existing models [6–9]. Additionally, acknowledging that KD affects children aged 7 to 10 years as well, we broadened our study's demographic scope to include newborns to 10-year-olds. By incorporating a larger number of cases, we aimed to enhance the quality of our Z-Score system and compare its normalization outcomes with those from previous studies.

Table 1 Age distribution of full training dataset

Age Group (years)	Patient Number
< 0.5	96
≥ 0.5, < 1	96
≥ 1, < 2	104
≥ 2, < 3	102
≥ 3, < 4	101
≥ 4, < 5	109
≥ 5, < 6	101
≥ 6, < 7	122
≥ 7, < 8	128
≥ 8, < 9	105
≥ 9, < 10	116

Material and methods

This section details the methodology and results of our study aimed at developing Z-Score models for normalizing measurements of three coronary arteries in children aged newborn to 10 years to enhance the diagnosis of Kawasaki disease (KD). The study included 1180 participants free from coronary or heart diseases. Measurements of the LCA, LAD, and RCA were collected, along with patient demographics such as age, height, weight, and body surface area (BSA). Several Z-Score models were proposed for different coronary arteries and age groups, with separate designs leading to multiple models. The models were rigorously tested for normality, and various data transformation techniques, including the Box-Cox method, were employed to optimize the RCA model. This section provides a comprehensive overview of the participant data, data transformation processes, basic and modified models, and the selection of our modeling efforts. Additionally, we further compared the predictive performance of each model. First, we calculated the Z-Scores for each age-specific model separately and then used these Z-Scores to compare the sensitivity and specificity across the entire dataset.

Participants

Data collection was conducted at Kaohsiung Chang Gung Memorial Hospital's Children's Medical Center (n = 831) and through a free clinic project in Kaohsiung's elementary schools (n = 349), encompassing a total of 1180 children. Participants were selected based on criteria that included being under 10 years of age, having undergone echocardiographic examinations, and having no history of coronary issues or congenital heart diseases. Collected data comprised the patients' age, height, weight, and BSA, along with measurements of the

internal lumen diameters of their LCA, LAD, and RCA. Table 1 presents the age distribution of the full training dataset. The mean age of the participants was 4.7 years. Due to variances in echocardiographic parameters, some LAD data were incompatible, resulting in 938 valid entries, while the LCA and RCA data both included 1180 entries. The LCA and LAD data passed the Anderson–Darling normality test, but the RCA data did not. On the other hand, the testing cohort comprised of 112 healthy subjects from a training dataset and 112 KD patients with coronary artery dilation, designated as the disease group. The mean age of the normal group and disease group was 4.80 and 4.07 years, separately. In the randomly selected testing dataset, the age distribution of the normal group closely resembled that of the full training dataset. Due to the nature of Kawasaki Disease, there are fewer individuals older than six years in the KD group (Table 2).

Data transformation

To establish predictive models, methods including linear, logarithmic, exponential, and square-root regression are employed. The key differences among these models lie in the BSA transformation and the incorporation of age variables. We evaluated various BSA calculation methods, including those proposed by Mosteller, DuBois, Haycock, and Gehan. Additionally, we reanalyzed several regression models from existing literature, recalculating their parameters and Z-Scores. In these models, M represents the measured coronary vessel diameter, while α and β are constants reflecting the impact of different parameters. The Anderson–Darling test revealed that none of the regression models from the literature could simultaneously normalize the Z-Score distributions for our LCA, LAD, and RCA data. Examining the distributions of the LCA, LAD, and RCA data in relation to BSA, we noted inhomogeneous variation in data distributions, irrespective of conditional distributions of LCA, LAD, or RCA data against varying BSA values.

The Box-Cox transformation (Eq. 1) [15], a widely used technique in data transformation, effectively minimizes the impact of inhomogeneous variation in regression models. The Box-Cox transformation of the variable M is defined as $M^{(\lambda)}$, where λ is the transformation parameter (Table 3).

We presented a simplified regression model (Eq. 2) using $M^{(\lambda)}$ as the dependent variable and BSA as the independent variable:

$$M^{\lambda} = \begin{cases} \frac{M^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(M) & \text{if } \lambda = 0 \end{cases} \quad (1)$$

$$M_i^{\lambda} = \alpha + \beta \times (BSA_i) + \varepsilon_i \quad (2)$$

In this model, each value in the dataset $(M_i, BSA_i)|i=1, \ldots n$ is transformed into a corresponding value in the set $(M_i^{(\lambda)}, BSA_i)|i=1, \ldots n$, fitting an optimal regression line. Assuming that the set ε_i, $i=1, \ldots n$ is normally distributed with a mean of 0 and a variance of σ^2, the optimal λ value is estimated using the maximum likelihood approach. The best regression model is then identified.

Adhering to Occam's razor, we prioritized models achieving normalization with minimal complexity. In evaluating models, both the fit of the Z-Score distribution to a normal distribution and a higher R^2 value are essential. For models with both normally distributed Z-Scores and R^2 values of ≥ 0.5, we also considered the calculation complexity and the number of independent variables.

Basic model

In our foundational model, $M_i^{(\lambda)}$ represents the ith data point for LCA, LAD, or RCA; BSA_i is the BSA value calculated from the patient's height and weight; and ε_i is an independently and normally distributed residual term with a mean of 0. The parameters λ, α, β, and γ in our optimal model are derived from our dataset. We calculated the Z-Score for each data point using the equation (Eqs. 3 and 4), where Z_i represents the Z-Score of the ith data point in the optimal regression model, and the standard error is the square root of the mean squared

Table 2 Age distribution of testing dataset

Age Group (years)	Normal Patient Number	KD Patient Number
< 0.5	3	8
≥ 0.5, < 1	11	1
≥ 1, < 2	11	15
≥ 2, < 3	9	16
≥ 3, < 4	14	14
≥ 4, < 5	10	26
≥ 5, < 6	10	11
≥ 6, < 7	16	5
≥ 7, < 8	8	2
≥ 8, < 9	9	10
≥ 9, < 10	11	4

Table 3 Box-cox Transformation

λ	$M^{(\lambda)}$	λ	$M^{(\lambda)}$
−2	$\frac{1}{M^2}$	0.5	$\sqrt{M}$
−1	$\frac{1}{M}$	1	M
−0.5	$\frac{1}{\sqrt{M}}$	2	M^2
0	ln(M)		

error. Equations 3 and 4 illustrate the final version of our selected model.

$$M_i^{(\lambda)} = \alpha + \beta \times ln\,(BSA_i) + \varepsilon_i \tag{3}$$

$$Z_i = \frac{M_i^{\lambda} - \alpha - \beta \times \ln(BSA_i)}{\sqrt{MSE}} \tag{4}$$

In our research, we considered the impact of age and other variations. However, they did not improve the model performance. For the equation and results about the candidate models, we show them in Appendix A and Appendix B.

Modified model for RCA data

Regarding RCA data, initial modeling attempts without the Box-Cox transformation yielded nonnormal Z-Scores. Although our application of the Box-Cox transformation decreased data variation inconsistency, the RCA Z-Scores persisted in being nonnormally distributed when not including LCA and LAD values as independent variables. Consequently, RCA characteristics may not be fully explained solely by variables like BSA, age, and sex. Due to the nonnormal distribution of the RCA data, we restructured the model to incorporate LCA and LAD parameters (Eqs. 5 and 6). This adjustment aimed to produce more accurate RCA results aligning with the data. The refined model is detailed below:

$$M_i^{(\lambda)} = \alpha + \beta_1 \times LCA + \beta_2 \times LAD + \beta_3 \times ln\,(BSA_I) + \varepsilon_I \tag{5}$$

$$Z_i = \frac{M_i^{\lambda} - \alpha - \beta_1 \times LCA - \beta_2 \times LAD - \beta_3 \times \ln(BSA_i)}{\sqrt{MSE}} \tag{6}$$

Age group selection

To manage the extensive combinations of age groups ranging from newborns to 10 years old, we implemented a systematic approach using a for-loop in R language to pair and test different age groupings. This allowed us to conduct a series of combi- natorial tests, assessing each combination against the Anderson–Darling normality test. For those models that successfully passed the normality test, we further evaluated them based on their R-square values, prioritizing models with higher R-square values for final selection. This approach ensured that the selected models were both statistically robust and provided the best fit for the data.

Results

Result of basic model

This section presents the result of the model performance of the Z-Score models for the three coronary arteries. To identify the optimal grouping for achieving normality, we test various combinations across different age groups. We ultimately found that combining ages 0 to 6 into a single group was effective in satisfying the normality assumption. However, for older age groups, maintaining one-year intervals was necessary to achieve normality. Therefore, there are five age groups include 0–6 years, 6–7 years, 7–8 years, 8–9 years, and 9–10 years. Table 4 displays the finalized parameters of all age models. Post Box-Cox transformation, both LCA and LAD data were successfully transformed into stable, normally distributed Z-Scores.

Table 4 Model parameters for different ages

Age Group		λ	α	β_1	β_2	β_3	MSE
0-6y	LCA	0.4646	1.2090	0.5581			0.0423
(n = 709)	LAD	0.1414	0.7369	0.3390			0.0321
	RCA	0.5859	0.0579	0.1587	0.2450	0.1970	0.0439
6-7y	LCA	0.6667	1.3051	0.8421			0.0910
(*n* = 122)	LAD	0.3030	0.8082	0.3966			0.0471
	RCA	0.5051	−0.2735	0.2375	0.2843	−0.0048	0.0406
7-8y	LCA	−0.2626	0.7947	0.1977			0.0186
(n = 128)	LAD	1.0303	1.0504	0.7779			0.1166
	RCA	0.5859	0.1061	0.0965	0.3106	0.2422	0.0440
8-9y	LCA	0.0202	0.9077	0.3947			0.0270
(*n* = 105)	LAD	0.9091	1.0298	0.8624			0.0794
	RCA	−0.7475	0.1695	0.0838	0.0969	0.0318	0.0048
9-10y	LCA	0.9899	1.4410	0.9222			0.1523
(*n* = 116)	LAD	1.1919	1.0901	1.2752			0.1639
	RCA	0.1414	0.0934	0.0109	0.3231	0.2721	0.0257

Table 5 Normality test & R^2 for different age models

Age Group		LCA	LAD	RCA
0-6y	Normality Test	0.792	0.251	0.600
	R^2	0.500	0.269	0.401
6-7y	Normality Test	0.102	0.064	0.211
	R^2	0.126	0.056	0.455
7-8y	Normality Test	0.088	0.449	0.008
	R^2	0.033	0.094	0.338
8-9y	Normality Test	0.459	0.284	0.396
	R^2	0.135	0.198	0.457
9-10y	Normality Test	0.214	0.774	0.068
	R^2	0.123	0.204	0.475

In Table 5, the adjusted R^2 for LAD diameter consistently showed low values across all models; enhancements such as increasing variable count (e.g., adding age or sex) or model complexity (e.g., transitioning to polynomial regression) did not elevate the R^2. This pattern suggests that the LAD diameter might possess less understood attributes. A higher R^2 value is indicative of a stronger linear relationship between variables, while a low R^2 value signals a potentially unsuitable model.

In line with Occam's razor, considering the models' R^2 values and Z-Score normality, we identified the optimal model for LCA and LAD data as a natural logarithmic model. Conversely, the RCA required a distinct approach, utilizing the modified model that integrates LCA and LAD values as independent variables.

Result of modified model

Incorporating LCA and LAD data into the modified model significantly enhanced both the Z-Score distribution and the adjusted R^2 for RCA diameters, leading to a more robust fit across all age groups. These improvements, as shown in Eq. 5, and the detailed results in Tables 4 and 5, demonstrate the effectiveness of our data-fitting methodology when compared to existing models. Specifically, our model performed well across various age groups in Table 6, passing normality tests and yielding competitive adjusted R^2 values for each coronary artery. The enhancements in our approach allowed for better handling of outliers and non-normal distributions, which are common challenges in coronary artery data.

The overall performance of our model, evaluated through ROC curve analysis (Fig. 1), showed AUC values of 0.975, 0.981, 0.979, and 0.979 for the Kuo, Lin et al. [7], Dallaire et al. [6], and McCrindle et al. [8] models, respectively. These results indicate that our model provides superior specificity (Table 7) while maintaining competitive sensitivity. Given the trade-off between sensitivity and specificity, our model demonstrated the best overall balance, outperforming the other models in terms of its ability to detect coronary artery dilation with a high degree of accuracy and reliability. This makes it a valuable tool for clinical decision-making in diagnosing Kawasaki Disease, especially when precision in coronary artery measurements is critical.

Discussion

Our study notably contributes to the existing body of literature by comparing newly developed Z-Score models with those from earlier research. We uniquely utilized Taiwanese data covering a broad age spectrum, from newborns to 10-year-old, and a substantially larger sample size. This approach allowed us to effectively tackle the

Table 6 Different age normality test & R^2 comparison with the existing models

Age Group	Methods	Anderson-Darling Test (p-value)			Adjusted R-Square		
		LCA	LAD	RCA	LCA	LAD	RCA
0-6y	Lin et al. [7]	0.253	0.217	0.011	0.514	0.272	0.424
	Dallaire et al. [6]	0.047	0.000	0.000	0.477	0.253	0.381
	McCrindle et al. [8]	0.253	0.217	0.011	0.514	0.272	0.424
	This paper	0.792	0.251	0.600	0.500	0.269	0.401
6-7y	Lin et al. [7]	-	-	-	-	-	-
	Dallaire et al. [6]	0.062	0.087	0.327	0.132	0.062	0.057
	McCrindle et al. [8]	0.028	0.022	0.540	0.118	0.054	0.052
	This paper	0.101	0.064	0.211	0.126	0.056	0.455
7-8y	Lin et al. [7]	-	-	-	-	-	-
	Dallaire et al. [6]	0.000	0.434	0.074	0.030	0.096	0.079
	McCrindle et al. [8]	0.030	0.033	0.065	0.033	0.082	0.060
	This paper	0.088	0.449	0.008	0.033	0.094	0.338
8-9y	Lin et al. [7]	-	-	-	-	-	-
	Dallaire et al. [6]	0.434	0.261	0.004	0.150	0.201	0.115
	McCrindle et al. [8]	0.453	0.034	0.137	0.134	0.177	0.103
	This paper	0.459	0.284	0.396	0.135	0.198	0.457
9-10y	Lin et al. [7]	-	-	-	-	-	-
	Dallaire et al. [6]	0.269	0.603	0.002	0.130	0.204	0.222
	McCrindle et al. [8]	0.246	0.017	0.538	0.110	0.183	0.220
	This paper	0.214	0.774	0.068	0.123	0.204	0.475

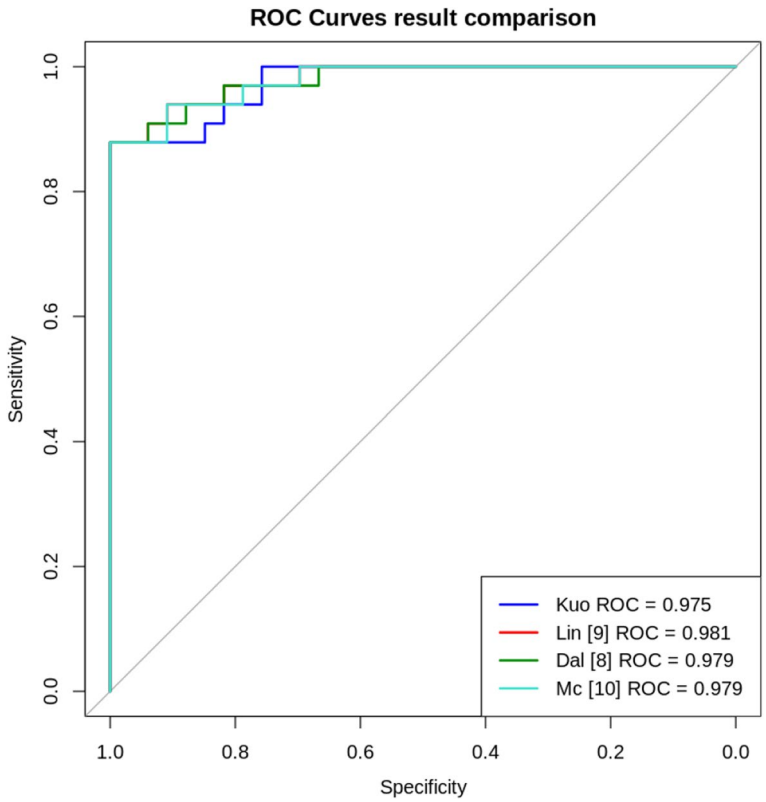


Fig. 1 ROC curve comparison

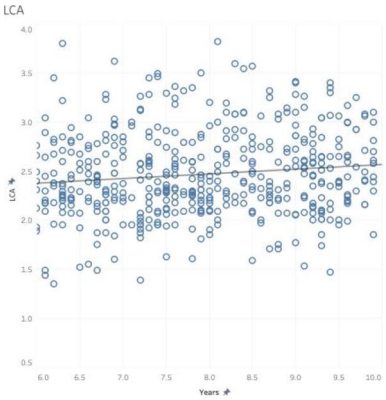
issue of unconventional data distribution which has been a challenge in previous studies. While the Taiwan Society of Pediatric Cardiology's 2014 study was limited to 412 children aged 0 to 6 years, our research expanded the age range up to 10 years and included 1180 participants, creating a dataset that is more representative of the wider pediatric population.

Z-Scores prove to be a more reliable method for predicting coronary artery lesions (CALs) in Kawasaki disease (KD) compared to arterial diameter measurements. According to the 1984 criteria set by the Japanese Ministry of Health and Welfare, coronary artery dilatation in children under 5 years and 5 years or older is defined as a maximum internal lumen diameter of at least 3 mm and 4 mm, respectively, along with any signs of local dilatation. Furthermore, a coronary artery aneurysm is characterized as a lumen measuring 4 to 8 mm or being 1.5 times larger than an adjacent segment, while a giant coronary artery aneurysm has a lumen size of 8 mm or more. Ae et al. have reported that Z-Score criteria are notably more sensitive in identifying coronary artery dilatations in KD than the Japanese criteria, although this heightened sensitivity was not apparent in cases of giant aneurysms [5]. In contrast to the Japanese criteria, Z-Score criteria consistently show greater sensitivity for detecting coronary artery dilatations across all ages and for identifying aneurysms in patients as young as one year old. The difference in detection sensitivity between the Z-Score and Japanese criteria stems from both the size of the coronary artery and the age of the patient. In our research, we accounted for age variations and utilized the Box-Cox transformation for the normalization of RCA diameters, which made the Z-Scores more statistically robust by ensuring consistency with the assumption of normality.

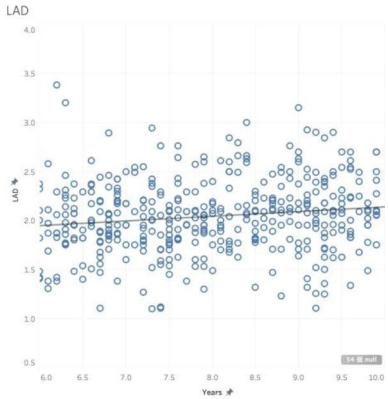
Table 7 Model performance comparison

	Sensitivity	Specificity
Lin et al. [7]	**0.697**	**1**
Dallaire et al. [6]	0.667	0.969
McCrindle et al. [8]	**0.697**	0.969
This paper	**0.697**	**1**

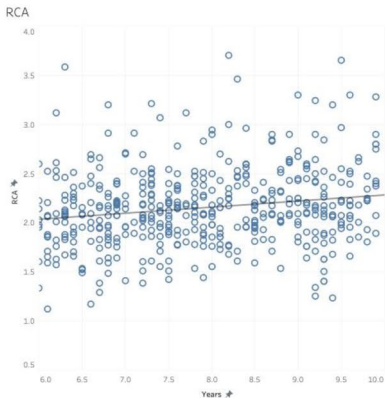
One common issue in previous studies is the inability to address the nonnormal distribution of arterial diameter data. In our study, we applied Box-Cox transformation to achieve more stable data, particularly for RCA diameters. Data transformed in this manner are capable of producing Z-Score results that meet the criteria



(a) Age & Diameter correlation of LCA. LCA=0.0477989*Years+2.08911. P-value=0.00787



(b) Age & Diameter correlation of LAD. LAD=0.0478745*Years+1.66031. P-value=0.00270

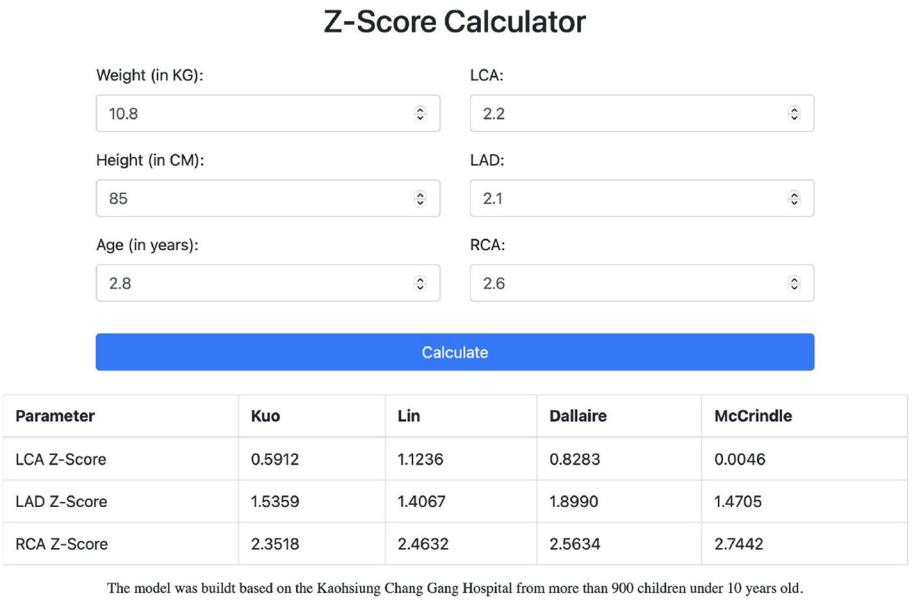


(c) Age & Diameter correlation of RCA. RCA=0.0619647*Years+1.66145. P-value=0.00013

Fig. 2 Age & coronary arteries diameter correlation

for normality. This technique can also be instrumental in developing Z-Score systems tailored to different populations.

When comparing the normal distribution of Z-Scores across different age groups, we cross-referenced various age combinations. The results showed a positive correlation



Parameter	Kuo	Lin	Dallaire	McCrindle
LCA Z-Score	0.5912	1.1236	0.8283	0.0046
LAD Z-Score	1.5359	1.4067	1.8990	1.4705
RCA Z-Score	2.3518	2.4632	2.5634	2.7442

The model was buildt based on the Kaohsiung Chang Gang Hospital from more than 900 children under 10 years old.

Fig. 3 Z-Score calculator

between coronary arteries diameter and body surface area for children under 6 years old. However, for those aged 6 and above, there was no correlation (Fig. 2). This lack of correlation might be the reason why data for children older than 6 years cannot be used to establish a unified model.

Even though the ROC value of the proposed model is slightly lower than existing models-by only 0.4% which is within the margin of random error, we believe that passing the normality test is essential in this study, especially before conducting Area Under the Curve (AUC) analysis. Zhou et al. [16] emphasize that understanding data distribution through normality testing directly affects the selection of appropriate analytical methods and the accuracy of confidence interval estimations. Ghasemi & Zahediasl [17] highlight that failing to meet normality assumptions can significantly impair statistical inference, leading to biased or invalid results. Hanley & McNeil [18] note that AUC calculations without verified distributional assumptions can yield misleading outcomes, potentially compromising the assessment of diagnostic accuracy. Therefore, neglecting normality testing can mask significant data quality issues, underscoring its importance for robust statistical validation in diagnostic research utilizing AUC analysis.

One limitation of this study is the complexity in understanding the data transformation methods and mathematical models, particularly the use of techniques like the Box-Cox transformation to fit the data to a normal distribution. While these approaches enhance the accuracy and robustness of the Z-Score models, they can be challenging for clinicians and researchers to interpret and apply in a practical setting, potentially limiting the accessibility and widespread adoption of these models in routine clinical practice. To overcome this issue, this study presents a web application [1] to resolve this problem. This system is shown in Fig. 3.

Conclusions

Our research has successfully developed the Kuo Z-Score models, derived from a collection of models that cater to a wide age range and encompass a substantial population size. By systematically testing various age group combinations through a programmatic approach and applying the Anderson–Darling normality test, we ensured the selection of robust models for children ranging from newborns to 10 years old. The application of the Box-Cox transformation further allowed us to normalize data distributions effectively. This enhanced Z-Score methodology provides clinicians with a more precise tool for identifying coronary artery dilatation and aneurysm in KD. The improved sensitivity and specificity rates highlight the model's clinical utility, offering significant advancements in KD diagnostic sensi- tivity and aiding in better clinical decision-making, potentially improving patient outcomes.

[1] https://worldstar.github.io/Z-ScoreCalculator/

Appendix A

Candidate Z-score models

When designing our optimal basic model, we conducted numerous variation models using BSA subjected to transformations i.e., logarithmic transformation in $Model_1$ (Eq. A1) and square-root transformation in $Model_3$ (Eq. A5). We also examined the effect of the age variable denoted as $year_i$ in $Model_1$. Candidate models are listed as follows:

$$Model_1 : M_i^{(\lambda)} = \alpha + \beta * ln(BSA_i) + \gamma * year_i + \epsilon_i \quad \text{(A1)}$$

$$Z - Score_1 : Z_i = \frac{M_i^{(\lambda)} - \alpha - \beta * ln(BSA_i) - \gamma * year_i}{Standard\ Error} \quad \text{(A2)}$$

$$Model_2 : M_i^{(\lambda)} = \alpha + \beta * BSA_i + \epsilon_i \quad \text{(A3)}$$

$$Z - Score_2 : Z_i = \frac{M_i^{(\lambda)} - \alpha - \beta * BSA_i}{Standard\ Error} \quad \text{(A4)}$$

$$Model_3 : M_i^{(\lambda)} = \alpha + \beta * \sqrt{BSA_i} + \epsilon_i \quad \text{(A5)}$$

$$Z - Score_3 := \frac{M_i^{(\lambda)} - \alpha - \beta * \sqrt{BSA_i}}{Standard\ Error} \quad \text{(A6)}$$

We also performed the same comparisons for the modified model:

$$Model_4 : M_i^{(\lambda)} = \alpha + \beta_1 * LCA + \beta_2 * LAD + \beta_3 * ln(BSA_i) + \gamma * year_i + \epsilon_i \quad \text{(A7)}$$

$$Z - Score_4 : Z_i = \frac{M_i^{(\lambda)} - \alpha - \beta_1 * LCA - \beta_2 * LAD - \beta_3 * ln(BSA_i) - \gamma * year_i}{Standard\ Error} \quad \text{(A8)}$$

$$Model_5 : M_i^{(\lambda)} = \alpha + \beta_1 * LCA + \beta_2 * LAD + \beta_3 * BSA_i + \epsilon_i \quad \text{(A9)}$$

$$Z - Score_5 : Z_i = \frac{M_i^{(\lambda)} - \alpha - \beta_1 * LCA - \beta_2 * LAD - \beta_3 * BSA_i}{Standard\ Error} \quad \text{(A10)}$$

$$Model_6 : M_i^{(\lambda)} = \alpha + \beta_1 * LCA + \beta_2 * LAD + \beta_3 * \sqrt{BSA_i} + \epsilon_i \quad \text{(A11)}$$

$$Z - Score_6 : Z_i \frac{M_i^{(\lambda)} - \alpha - \beta_1 * LCA - \beta_2 * LAD - \beta_3 * \sqrt{BSA_i}}{Standard\ Error} \quad \text{(A12)}$$

Appendix B

Additional normality test and R^2 comparison of candidate models

The full results of the normality test for the basic models are listed in Tables 8 and 9. The dataset was conducted as one group here. The RCA Z-Scores of the models did not pass the normality test.

Table 8 Full normality test & R^2 result of basic model for LCA and LAD data

Model type	Normality Test			R^2		
	LCA	LAD	RCA	LCA	LAD	RCA
Model 1	0.425	0.349	0.001	0.500	0.267	0.409
Model 2	0.134	0.402	0.003	0.466	0.258	0.379
Model 3	0.552	0.314	0.001	0.489	0.267	0.399

Table 9 Normality Test and R^2 of Modified Model for RCA Data

Model type	Normality Test	R^2
	RCA	RCA
Model 4	0.572	0.401
Model 5	0.546	0.395
Model 6	0.608	0.398

The R^2 results indicated minimal variation among the models. Our selected model slightly outperformed $Z - Score_1$ (Eq. A2) in the normality test.

Abbreviations

LCA Left main Coronary Artery
LAD Left Anterior Descending coronary artery
RCA Right Coronary Artery

Acknowledgements

This study gratefully acknowledges the support received from the National Science and Technology Council and Kaohsiung Chang Gung Memorial Hospital.

Authors' contributions

H.C. Kuo provided ultrasound images and supervised the study. The data set is ac- cessible upon request. Please contact Prof. H.C. Kuo (email: erickuo48@yahoo.com.tw) to obtain further information. Y.H. Chen contributed to the study design. Y.C. Lin annotated medical images and wrote the paper. W.Y. Cheng wrote the R language code and conducted experiments. Y.F. Chen performed statistical analysis. S.F. Liu and M.H. Guo conducted most echocardiography imaging during the investiga- tion. S.H. Chen is the paper's guarantor.

All authors contributed to the manuscript revision, and read, and approved the submitted version.

Funding
This work was supported by the National Science and Technology Council [project numbers: NSTC 113-2221-E-032-022, NSTC 113-2221-E-182-064, NSTC 112–2314-B-182–032-MY3, NSTC 109–2221-E-032–042 and NSTC 109–2221-E-230–006]; and Kaohsiung Chang Gung Memorial Hospital [project num- bers: CMRPD3M0011, CMRPD3N0011, CMRPD3P0011 on study design and data collection and analysis, CORPG8N0071, COPRG8M0091, CORPG8P0021, CORPG8L0321 on study design and CMRPG8M1321 on interpretation of data].

Data availability
The datasets analyzed during the current study are not publicly available due to privacy protection regulations of Chang Gung Memorial Hospital but are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
The echocardiographic images utilized in this study were obtained from Chang Gung Memorial Hospital. The study was conducted in accordance with the guidelines set by the Institutional Review Board (IRB) of the hospital and received approval with the IRB number No. 202101666B0.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Pediatrics, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan. [2]College of Medicine, Chang Gung University, Taoyuan, Taiwan. [3]Kawasaki Disease Center, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan. [4]Department of Respiratory Therapy, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan. [5]Department of Computer Science and Information Engineering, Tamkang University, New Taipei City, Taiwan. [6]Department of Management Sciences, Tamkang University, New Taipei City, Taiwan. [7]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taoyuan, Taiwan. [8]Department of Information Management, Chang Gung University, Taoyuan, Taiwan.

Received: 3 July 2024 Accepted: 28 November 2024
Published online: 21 December 2024

References

1. Darby J, Jackson J. Kawasaki Disease and Multisystem Inflammatory Syndrome in Children: An Overview and Comparison. Am Fam Physician. 2021;104(3):244–52.
2. Altammar F, Lang B. Kawasaki disease in the neonate: case report and literature review. Pediatr Rheumatol. 2018;16:1–6.
3. Yang TJ, Lin MT, Lu CY, Chen JM, Lee PI, Huang LM, et al. The prevention of coronary arterial abnormalities in Kawasaki disease: A meta-analysis of the corticosteroid effectiveness. J Microbiol Immunol Infect. 2018;51(3):321–31.
4. Bratincsak A, Reddy VD, Purohit PJ, Tremoulet AH, Molkara DP, Frazer JR, et al. Coronary artery dilation in acute Kawasaki disease and acute illnesses associated with Fever. Pediatr Infect Dis J. 2012;31(9):924–6.
5. Ae R, Shibata Y, Kobayashi T, Kosami K, Kuwabara M, Makino N, et al. Differences in Sensitivity Between the Japanese and Z Score Criteria for Detecting Coronary Artery Abnormalities Resulting from Kawasaki Disease. Pediatr Cardiol. 2022;44(1):153–60.
6. Dallaire F, Dahdah N. New equations and a critical appraisal of coronary artery Z scores in healthy children. J Am Soc Echocardiogr. 2011;24(1):60–74.
7. Lin MT, Chang CH, Hsieh WC, Chang CE, Chang YM, Chen YC, et al. Coronary diameters in Taiwanese children younger than 6 years old: Z-score regression equations derived from body surface area. Acta Cardiol Sin. 2014;30(4):266–73.
8. McCrindle BW, Li JS, Minich LL, Colan SD, Atz AM, Takahashi M, et al. Coronary artery involvement in children with Kawasaki disease: risk factors from analysis of serial normalized measurements. Circulation. 2007;116(2):174–9.
9. Olivieri L, Arling B, Friberg M, Sable C. Coronary artery Z score regression equations and calculators derived from a large heterogeneous population of children undergoing echocardiography. J Am Soc Echocardiogr. 2009;22(2):159–64.
10. Kobayashi T, Fuse S, Sakamoto N, Mikami M, Ogawa S, Hamaoka K, et al. A New Z Score Curve of the Coronary Arterial Internal Diameter Using the Lambda-Mu-Sigma Method in a Pediatric Population. J Am Soc Echocardiogr. 2016;29(8):794-801.e29.
11. Lopez L, Colan S, Stylianou M, Granger S, Trachtenberg F, Frommelt P, et al. Relationship of Echocardiographic Z Scores Adjusted for Body Surface Area to Age, Sex, Race, and Ethnicity. Circ Cardiovasc Imaging. 2017;10(11):e006979.
12. Gyeong-Hee Y. Characteristics of z score systems for diagnosing coronary abnormalities in Kawasaki disease. Clin Exp Pediatr. 2022;65(9):448–9.
13. Hye KS. Diagnosis of coronary artery abnormalities in Kawasaki disease: recent guidelines and z score systems. Clin Exp Pediatr. 2022;65(9):430–8.
14. Lorenzoni RP, Elkins N, Quezada M, Silver EJ, Mahgerefteh J, Hsu DT, et al. Impact of Z score system on the management of coronary artery lesions in Kawasaki disease. Cardiol Young. 2022;32(6):952–9.
15. Chatfield C. The analysis of time series: An introduction. Chapman and hall/CRC; 2003
16. Zhou XH, Obuchowski NA, McClish DK. Statistical Methods in Diagnostic Medicine. 2nd ed. Wiley Series in Probability and Statistics. John Wiley & Sons; 2014
17. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: A guide for non-statisticians. International Journal of Endocrinology and Metabolism. 2012;10(2):486–9.
18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.

Publisher's Note